

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:34 ; Search time 3675.59 Seconds
(without alignments)
8198.470 Million cell updates/sec

Title: US-09-497-957-9
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAGATTTCCTCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_or.*

21: em_ov.*

22: em_pat.*

23: em_ph.*

24: em_pi.*

25: em_ro.*

26: em_sts.*

27: em_un.*

28: em_vl.*

29: em_hgt_hum.*

30: em_hgt_inv.*

31: em_hgt_inv.*

32: em_hgt_inv.*

33: em_hgt_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query
No. Score Match Length DB ID Description

1	1440	100.0	1440	6	AR117793	Sequence
2	1440	100.0	1440	6	AR149463	Sequence
3	1440	100.0	2727	9	HSU60319	Homo sapien
4	1438.4	99.9	1440	6	AR117794	Sequence
5	1438.4	99.9	1440	6	AR117795	Sequence
6	1438.4	99.9	1440	6	AR149464	Sequence
7	1438.4	99.9	1440	6	AR149465	Sequence
8	1436.8	99.8	1440	6	AR117796	Sequence
9	1436.8	99.8	1440	6	AR149466	Sequence
10	1201	83.4	1280	9	HS249335	Homo sapi
11	1189	82.6	1200	9	AF115265	Homo sapi
12	993	62.0	1045	9	AF079407	Homo sapi
13	897	62.3	1320	4	AY007541	Ceratotheri
14	897	62.3	1320	4	AY007543	Dicerorhi
15	896.8	62.3	1319	4	AY007544	Rhinocero
16	893.8	62.1	2332	4	AY007542	Dicerorhi
17	849.4	59.0	1085	9	HS249336	Homo sapi
18	787	54.7	1073	9	HS249337	Homo sapi
19	775.2	53.8	1885	9	AF144242	Homo sapi
20	711.4	49.4	823	9	AF079408	Homo sapi
21	637.6	44.3	1706	10	RNHERHAEM	AJ001517 Rattus no
22	622.8	43.2	1529	10	MMU66849	Mus musculu
23	617.4	42.9	781	9	AF079409	Homo sapi
24	573	39.8	809	9	HS250635	Homo sapi
25	549.8	38.2	987	9	AF150664	Homo sapi
26	486.8	33.8	686	9	AF144240	Homo sapi
27	478.2	33.2	789	10	AF008587	Rattus no
28	464.8	32.3	804	9	AF149804	Homo sapi
29	448.8	31.2	819	10	AF176534	Homo sapi
30	436.4	30.3	735	9	AF115264	Homo sapi
31	393	27.3	560	9	AF144243	Homo sapi
32	322	22.4	874	9	HS249337	Homo sapi
33	321	22.3	10825	6	AR117789	Sequence
34	321	22.3	10825	6	AR117790	Sequence
35	321	22.3	10825	6	AR149459	Sequence
36	321	22.3	10825	6	AR149460	Sequence
37	321	22.3	12146	9	HS249337	Sequence
38	321	22.3	193752	2	AL359892	Homo sapi
39	321	22.3	246240	6	AR036572	Sequence
40	321	22.3	246240	6	AR036573	Sequence
41	321	22.3	246240	6	AR036574	Sequence
42	321	22.3	246282	9	HS091328	Human hered
43	320.4	22.2	492	9	AF144241	Homo sapi
44	319.4	22.2	10825	6	AR117791	Sequence
45	319.4	22.2	10825	6	AR117792	Sequence

ALIGNMENTS

RESULT	1	AR117793	Sequence	9	from patent US 6140305.	DNA	Linear	PAT 16-MAY-2001
LOCUS	AR117793	1440 bp						
DEFINITION	Sequence	9	from patent US 6140305.					
ACCESSION	AR117793							
VERSION	AR117793.1	GI:14098699						
KEYWORDS	Unknown.							
SOURCE	Unknown.							
ORGANISM	Unknown.							

REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas.W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 9 31-OCT-2000;
FEATURES Location/Qualifiers
source 1. 1440
/organism="unknown"

BASE COUNT 347 a 355 c 407 g 331 t
ORIGIN

Query Match 100.0%; Score 1440; DB 6; Length 1440;

Best Local Similarity 100.0%; Pred. No. 0;				
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps				
QY	1	GGGACACATGGATACACTAGTGTGTTTCACAAAGCAGACTACTTCTGCTGTAGGAGAGAGA	60	
Db	1	GGGACACATGGATACACTAGTGTGTTTCACAAAGCAGACTACTTCTGCTGTAGGAGAGAGA	60	
QY	61	ACTAAAGTTCTGAAGACCTGTGCTTCTCACCAGGAAGTTTACTGGGCATCTCCTGAG	120	
Db	61	ACTAAAGTTCTGAAGACCTGTGCTTCTCACCAGGAAGTTTACTGGGCATCTCCTGAG	120	
QY	121	CCTAGGCAATAGCTGTAGGCTGACTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG	180	
Db	121	CCTAGGCAATAGCTGTAGGCTGACTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG	180	
QY	181	AAGCGAGATTTAAGCGGACGTGGCGGCAGAGCTGGGGAAATGGGCCCGGAGCCAGGC	240	
Db	181	AAGCGAGATTTAAGCGGACGTGGCGGCAGAGCTGGGGAAATGGGCCCGGAGCCAGGC	240	
QY	241	CGCGCGTTCTCCTCCCTGATGCTTTTGCAGACCGCGGCTCTGCAGGGCGCTTGCCTGGGTT	300	
Db	241	CGCGCGTTCTCCTCCCTGATGCTTTTGCAGACCGCGGCTCTGCAGGGCGCTTGCCTGGGTT	300	
QY	301	CACACTCTGTGCACTACCTCTTCATGGGTGCCTCAGACGAGACCTTGGTCTTTCTCTGT	360	
Db	301	CACACTCTGTGCACTACCTCTTCATGGGTGCCTCAGACGAGACCTTGGTCTTTCTCTGT	360	
QY	361	TTGAAGCTTTGGGCTACGTGGATGACCACTGTTCCGTGTTCTATGATCATGAGAGTCGCC	420	
Db	361	TTGAAGCTTTGGGCTACGTGGATGACCACTGTTCCGTGTTCTATGATCATGAGAGTCGCC	420	
QY	421	GTGTGGAGCCCGAAGTCCATGGTTTCCAGTAGAAATTTCAAGCCAGATGTGGCTGCAGC	480	
Db	421	GTGTGGAGCCCGAAGTCCATGGTTTCCAGTAGAAATTTCAAGCCAGATGTGGCTGCAGC	480	
QY	481	TGAGTCAGAGTCTGAAAGGGTGGGATCATGTTCCACTGTTGACTCTTGGACTATTATGG	540	
Db	481	TGAGTCAGAGTCTGAAAGGGTGGGATCATGTTCCACTGTTGACTCTTGGACTATTATGG	540	
QY	541	AAATCACAACACAGCAAGAGAGTCCACACCCTCGAGGTCATCCTGGGCTGTCAATATGC	600	
Db	541	AAATCACAACACAGCAAGAGAGTCCACACCCTCGAGGTCATCCTGGGCTGTGAAATATGC	600	
QY	601	AAGAAGCAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCGAGGACCACTTG	660	
Db	601	AAGAAGCAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCGAGGACCACTTG	660	
QY	661	AATTCTGCCCTGCACACTTGGATTGGAGAGCAGCAAGAACCCAGGGCCTGGCCCAACCAAGC	720	
Db	661	AATTCTGCCCTGCACACTTGGATTGGAGAGCAGCAAGAACCCAGGGCCTGGCCCAACCAAGC	720	
QY	721	TGAGTGTGGNAAGGCACAAGATTGGGGCCAGCACAAGGGCTACTTGGAGAGGAGACT	780	
Db	721	TGAGTGTGGNAAGGCACAAGATTGGGGCCAGCACAAGGGCTACTTGGAGAGGAGACT	780	
QY	781	GCCTTGACAGCTGACAGCTTGTCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGTGC	840	
Db	781	GCCTTGACAGCTGACAGCTTGTCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGTGC	840	
QY	841	CTCTTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG	900	
Db	841	CTCTTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG	900	
QY	901	CCTTGAAGTACTACCCCCAGAATCATCACCATTGAAGTGGCTGAAGGATTAAGCAGCAATGG	960	
Db	901	CCTTGAAGTACTACCCCCAGAATCATCACCATTGAAGTGGCTGAAGGATTAAGCAGCAATGG	960	
QY	961	ATGCCAAGGATTCGAACCTTAAGACGTATTGCCCAATGGGGATGGGACCTTACCAAGGCT	1020	
Db	961	ATGCCAAGGATTCGAACCTTAAGACGTATTGCCCAATGGGGATGGGACCTTACCAAGGCT	1020	
QY	1021	GGATAACCTTGGCTGTACCCCTGGGGGAAGACGAGAGATATTACGTGCCAGGTGGAGCACC	1080	

D	b	1021	GGAATAACCTTTGGCTGTACCCCTCGGGAAGACAGAGATATACGTGCCAGGTGGAGCAC	1080
Q	y	1081	CAGCCTTGATCATAGCCCTCATTTGTGATCTGGAGCCCTCAACCCTCTGGCACCTTAGTCA	1140
D	b	1081		
D	b	1081	CAGCCTTGATCATAGCCCTCATTTGTGATCTGGAGCCCTCAACCCTCTGGCACCTAGTCA	1140
Q	y	1141	TGGAGTCATCAGTGGAAATTGCTGTTTTTGTGCGTCATCTTGTTTCATTGGAAATTTGTTCA	1200
D	b	1141		
D	b	1141	TGGAGTCATCAGTGGAAATTGCTGTTTTTGTGCGTCATCTTGTTTCATTGGAAATTTGTTCA	1200
Q	y	1201	TAATATTAAAGAAAGAGCAGGGTTCAAGAGGAGCCATGGGCACACTAGCTCTTAGCTGAAC	1260
D	b	1201		
D	b	1201	TAATATTAAAGAAAGAGCAGGGTTCAAGAGGAGCCATGGGCACACTAGCTCTTAGCTGAAC	1260
Q	y	1261	GTAGTGACACGAGCGCTCGAGACACACACACACACACACACACACACACACACACACAC	1320
D	b	1261		
D	b	1261	GTAGTGACACGAGCGCTCGAGACACACACACACACACACACACACACACACACACACAC	1320
Q	y	1321	AGGAGTGCATTTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAANA	1380
D	b	1321		
D	b	1321	AGGAGTGCATTTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAANA	1380
Q	y	1381	TGCTGACGACACCTTGATTTAGCCTTCTGTTTCATTTCTTCAAAGAAATTTCCCCA	1440
D	b	1381		
D	b	1381	TGCTGACGACACCTTGATTTAGCCTTCTGTTTCATTTCTTCAAAGAAATTTCCCCA	1440
<hr/>				
RESULT 2				
LOCUS AR149463 1440 bp DNA linear PAT 08-AUG-2001				
DEFINITION Sequence 9 from patent US 6228594.				
ACCESSION AR149463				
VERSION AR149463.1 GI:15114054				
KEYWORDS .				
SOURCE Unknown.				
ORGANISM Unknown.				
REFERENCE Unclassified.				
AUTHORS 1 (bases 1 to 1440)				
TITLE Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.				
METHOD Method for determining the presence or absence of a hereditary hemochromatosis gene mutation				
JOURNAL Patent: US 6228594-A 9 08-MAY-2001;				
FEATURES Location/Qualifiers				
source 1..1440				
BASE COUNT 347 a 355 c 407 g 331 t				
ORIGIN /organism="unknown"				
<hr/>				
Query Match 100.0%; Score 1440; DB 6; Length 1440;				
Best Local Similarity 100.0%; Pred.No. 0;				
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
<hr/>				
Q	y	1	GGGACACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCTGCTAGGAGAGAGA 60	
D	b	1	GGGACACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCTGCTAGGAGAGAGA 60	
Q	y	61	ACTAAAGTCTGAAAGACCTGTGCTTTTCACGAGGAGTTTACTGGGCATCTCTGAG 120	
D	b	61	ACTAAAGTCTGAAAGACCTGTGCTTTTCACGAGGAGTTTACTGGGCATCTCTGAG 120	
Q	y	121	CCTAGGCAATAGCTGTAGGCTGACTCTTGGAGCCATCCCGTTTCCCGCCCCCAAAAAG 180	
D	b	121	CCTAGGCAATAGCTGTAGGCTGACTCTTGGAGCCATCCCGTTTCCCGCCCCCAAAAAG 180	
Q	y	181	AAGCGGAGATTAAACGGGACCTCGGGCCAGAGCTGGGGAATGGGCCCGCAGCCAGGC 240	
D	b	181	AAGCGGAGATTAAACGGGACCTCGGGCCAGAGCTGGGGAATGGGCCCGCAGCCAGGC 240	
Q	y	241	CGGCGCTTCTCCTCATGCTCTTTGCAGACCGCGGTCCTGCAGGGGCGCTTCTCGGTT 300	
D	b	241	CGGCGCTTCTCCTCATGCTCTTTGCAGACCGCGGTCCTGCAGGGGCGCTTCTCGGTT 300	

QY 301 CACACTCTGCACTACCTCTTCAATGGGTGCCTCAGACGAGACCTTGGCTCTTCTTCTGT 360
Db 301 CACACTCTGCACTACCTCTTCAATGGGTGCCTCAGACGAGACCTTGGCTCTTCTTCTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTCGTCTTATGATCATGAGATCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTCGTCTTATGATCATGAGATCGCC 420
QY 421 GTGTGAGCCCGAAGTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGCGTCAGC 480
Db 421 GTGTGAGCCCGAAGTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGCGTCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 AAAATCACAAACACAGCAGAGGAGTCCACACCTTCAGGTCACTCGGCTGTGAAATGC 600
Db 541 AAAATCACAAACACAGCAGAGGAGTCCACACCTTCAGGTCACTCGGCTGTGAAATGC 600
QY 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGTATGATGGGAGGACCACTTG 660
Db 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGTATGATGGGAGGACCACTTG 660
QY 661 AATTCTGCCCTCAGACACTGGATTGGAGACGACAGACCCAGGCGCTGGCCCAACAGC 720
Db 661 AATTCTGCCCTCAGACACTGGATTGGAGACGACAGACCCAGGCGCTGGCCCAACAGC 720
QY 721 TGGAGTGGAAAGGACAGATTTGGGCGAGGCAGACAGGSCCTACCTGGAGAGGAGT 780
Db 721 TGGAGTGGAAAGGACAGATTTGGGCGAGGCAGACAGGSCCTACCTGGAGAGGAGT 780
QY 781 GCCTTCACAGCTGACAGAGTTCGTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 GCCTTCACAGCTGACAGAGTTCGTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGTGACACATCATGTGACCTCTTCAGTGACCACTCTAGGTGTCGGG 900
Db 841 CTCCTTTGGTGAAGTGACACATCATGTGACCTCTTCAGTGACCACTCTAGGTGTCGGG 900
QY 901 CTTGAACCTACTACCCCGAGAACATCACCATGAAGTGCCTGAAGGATGAAGCAGCAATGG 960
Db 901 CTTGAACCTACTACCCCGAGAACATCACCATGAAGTGCCTGAAGGATGAAGCAGCAATGG 960
QY 961 ATGCCAAGAGTTCGAACCTTAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGAGTTCGAACCTTAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACGACAGATATACGTGCCAGGTGGAGCAC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACGACAGATATACGTGCCAGGTGGAGCAC 1080
QY 1081 CAGGCTCGATCAGCCCTCATTTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA 1140
Db 1081 CAGGCTCGATCAGCCCTCATTTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA 1140
QY 1141 TTGGAGTCACTAGTGAATTCCTGTTTTCGTCATCTTGTTCATGGAATTTGTTCA 1200
Db 1141 TTGGAGTCACTAGTGAATTCCTGTTTTCGTCATCTTGTTCATGGAATTTGTTCA 1200
QY 1201 TAATATTAAAGGAAGGAGGCTTCAAGAGGAGCCATGGGCGACTACGTCTAGCTGAAC 1260
Db 1201 TAATATTAAAGGAAGGAGGCTTCAAGAGGAGCCATGGGCGACTACGTCTAGCTGAAC 1260
QY 1261 GTGAGTCACAGCCCTGACAGTCACTGTGGGAAGGACAGCAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTCACAGCCCTGACAGTCACTGTGGGAAGGACAGCAAACTAGAGACTCAAG 1320
QY 1321 AGGAGTGCATTTATGAGCTTCTCATGTTTCAGGAGAGGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTTCTCATGTTTCAGGAGAGGTTGAACCTAAACATAGAAAT 1380

QY 1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTCTCATTTCTCTCAAAAAGATTTCCCA 1440
Db 1381 TGCCTGACGAACTCCTTGATTTTAGCCTTCTCTCTCATTTCTCTCAAAAAGATTTCCCA 1440
RESULT 3
HSU60319 2727 bp mRNA linear PRI 29-OCT-1997
LOCUS Homo sapiens haemochromatosis protein (HLA-H) mRNA, complete cds.
DEFINITION
ACCESSION U60319
VERSION U60319.1 GI:1469789
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2727)
AUTHORS Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A.,
Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A.,
Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P.,
Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C.,
Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E.,
Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J.,
Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
A novel MHC class I-like gene is mutated in patients with
hereditary haemochromatosis
Nature Genet. 13 (4), 399-408 (1996)
96331279
REFERENCE 2 (bases 1 to 2727)
AUTHORS Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A.,
Basava,A., Dormishian,F., Domingo,R., Ellis,M.C., Fullan,A.,
Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P.,
Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C.,
Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E.,
Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J.,
Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.
Direct Submission
Submitted (10-JUN-1996) Mercator Genetics, 4040 Campbell Ave.,
Menlo Park, CA 94025, USA
FEATURES
Location/Qualifiers
1..2727
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
1..2727
/gene="HLA-H"
/note="HFE"
222..1268
/gene="HLA-H"
/codon_start=1
/product="haemochromatosis protein"
/protein_id="AAC51823.1"
/db_xref="GI:1469790"
/translation="MGPRAPALLMLLQTLAVLQGLRLRLSHSLHYLFMGASEDILGL
SLFEALGYVDOLFVFDHESRRVPTWSSRISSQMWLQSLQKGBHMTVD
WTIMENHNHSHKSTIQVTLGCEMEDNSTGYWKYGDQDLHFCPTDLDRAAEP
RAWPKLEWERHKIRARONRAYLERDCAQQLQELLEGRVLDQCVPLVKVTHVTS
SVTLTLCALNTYPTONITMKLWKDKPMDAKEFEKPKVLPNGDGYQSWITLAVPPGE
EORYTCQVEHPCLDQPLIVWEPSPSGTLVIGVISGIAVFVVILFGLFILRLRRQG
SRGAMGHVLAERE"
BASE COUNT 702 a 606 c 660 g 759 t
ORIGIN
Query Match 100.0%; Score 1440; DB 9; Length 2727;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGAGTACCTTCTGCTGTAGGAGAGAGA 60
Db 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGAGTACCTTCTGCTGTAGGAGAGAGA 60

QY 61 ACTAAGTTCTGAAGACCTGTGCTTTTCCACGAGAGTGTCTTACTGGGCATCTCCTGAG 120
|||||
Db 61 ACTAAGTTCTGAAGACCTGTGCTTTTCCACGAGAGTGTCTTACTGGGCATCTCCTGAG 120
|||||
QY 121 CTTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGCTTTCCCGCGCCCCAAAAG 180
|||||
Db 121 CTTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGCTTTCCCGCGCCCCAAAAG 180
|||||
QY 181 AAGCGGAGATTAAAGCGGACCTGGCGGCAGAGCTGGGGAATGGCGCGCGAGCCAGGC 240
|||||
Db 181 AAGCGGAGATTAAAGCGGACCTGGCGGCAGAGCTGGGGAATGGCGCGCGAGCCAGGC 240
|||||
QY 241 CGCGCTCTCTCTCTGATGCTTTTGCAGACCGCGTCTCGAGGGCGCTTGTGCGGT 300
|||||
Db 241 CGCGCTCTCTCTCTGATGCTTTTGCAGACCGCGGTCTCGAGGGCGCTTGTGCGGT 300
|||||
QY 301 CACACTCTCTGCACTACCTCTTTCATGGGTCCCTCAGAGCAGACCTTGGTCTTTCTTGT 360
|||||
Db 301 CACACTCTCTGCACTACCTCTTTCATGGGTCCCTCAGAGCAGACCTTGGTCTTTCTTGT 360
|||||
QY 361 TTGAAGCTTTGGGCTACGTTGGATGACAGCTGTCTGTCTTCTATCATGATGAGAGTCGCC 420
|||||
Db 361 TTGAAGCTTTGGGCTACGTTGGATGACAGCTGTCTGTCTTCTATCATGATGAGAGTCGCC 420
|||||
QY 421 GTGTGAGCGCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
|||||
Db 421 GTGTGAGCGCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
|||||
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCTACTGTTGACTCTCGGACTATTATGG 540
|||||
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCTACTGTTGACTCTCGGACTATTATGG 540
|||||
QY 541 AAAATCACAAACACAGAGAGTCCACACCTCGAGGTCTATCTCTGGGCTGTGAATGC 600
|||||
Db 541 AAAATCACAAACACAGAGAGTCCACACCTCGAGGTCTATCTCTGGGCTGTGAATGC 600
|||||
QY 601 AAGAAGAACACTACCGAGGGTCTTGGAGTACGGGTATGATGGGCGAGGACCACTTG 660
|||||
Db 601 AAGAAGAACACTACCGAGGGTCTTGGAGTACGGGTATGATGGGCGAGGACCACTTG 660
|||||
QY 661 AATTCCTGCCCTGACACATCATGTGATGGAGAGCAGACACCCAGGGCTCGGCCACCAAGC 720
|||||
Db 661 AATTCCTGCCCTGACACATCATGTGATGGAGAGCAGACACCCAGGGCTCGGCCACCAAGC 720
|||||
QY 721 TGGAGTGGGAAAGGCAACAAGATTCCGGCCAGGCGAGCAACAGGGCTACCTGGAGAGGACT 780
|||||
Db 721 TGGAGTGGGAAAGGCAACAAGATTCCGGCCAGGCGAGCAACAGGGCTACCTGGAGAGGACT 780
|||||
QY 781 GCCTCTGACAGCTGCAGCAGTTCTGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
|||||
Db 781 GCCTCTGACAGCTGCAGCAGTTCTGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
|||||
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTACGGTGTGCGG 900
|||||
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTACGGTGTGCGG 900
|||||
QY 901 CTTTGAACCTACTACCCCGCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGG 960
|||||
Db 901 CTTTGAACCTACTACCCCGCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGG 960
|||||
QY 961 ATGCCAAGGAGTTGGAACCTTAAGACCTATTGCCAATGGGATGGAGCTACCGAGGCT 1020
|||||
Db 961 ATGCCAAGGAGTTGGAACCTTAAGACCTATTGCCAATGGGATGGAGCTACCGAGGCT 1020
|||||
QY 1021 GGATAACCTTTGGCTGTATCCCGCTGGGGAAGAGAGATATACGTGCCAGGTGGAGCACC 1080
|||||
Db 1021 GGATAACCTTTGGCTGTATCCCGCTGGGGAAGAGAGATATACGTGCCAGGTGGAGCACC 1080
|||||
QY 1081 CAGGCTTGGATACGCCCTCATTTGTGATCTGGGAGGCCCTCACCGTCTGGCACCCCTAGTCA 1140
|||||
Db 1081 CAGGCTTGGATACGCCCTCATTTGTGATCTGGGAGGCCCTCACCGTCTGGCACCCCTAGTCA 1140
|||||
QY 1141 TTGAGTCACTAGTGGAAATGCTGTTTTTCTCGTCACTCTTCTCAATTTGGTTCAT 1200
|||||

Db 1141 TTGAGTCACTAGTGGAAATGCTGTTTTTCTCGTCACTCTTCTCAATTTGGAAATTTGTTC 1200
|||||
QY 1201 TAATATTAAAGAGAGCAGGTTCAAGAGAGCCATGGGGCACTAGCTCTTAGCTGAAC 1260
|||||
Db 1201 TAATATTAAAGAGAGCAGGTTCAAGAGAGCCATGGGGCACTAGCTCTTAGCTGAAC 1260
|||||
QY 1261 GTGAGTGACACGAGCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
|||||
Db 1261 GTGAGTGACACGAGCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
|||||
QY 1321 AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
|||||
Db 1321 AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
|||||
QY 1381 TGCTGTGAGCACTCTTCATGTTTATGAGCTCTCTCTTCATTTCTCAAAAAGATTTCCCCA 1440
|||||
Db 1381 TGCTGTGAGCACTCTTCATGTTTATGAGCTCTCTCTTCATTTCTCAAAAAGATTTCCCCA 1440
|||||
RESULT 4
AR117794 LOCUS AR117794 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 10 from patent US 6140305.
ACCESSION AR117794
VERSION AR117794.1 GI:14098700
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuchihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 10 31-OCT-2000;
FEATURES
Location/Qualifiers
Source 1..1440
BASE COUNT 348 a 355 c 406 g 331 t
ORIGIN
Query Match 99.9%; Score 1438.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGGACACTGGATCACTACCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
|||||
Db 1 GGGGACACTGGATCACTACCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
|||||
QY 61 ACTAAGTTCTGAAGACCTGTGCTTTTCCACGAGAGTGTCTTACTGGGCATCTCCTGAG 120
|||||
Db 61 ACTAAGTTCTGAAGACCTGTGCTTTTCCACGAGAGTGTCTTACTGGGCATCTCCTGAG 120
|||||
QY 121 CTTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGCTTTCCCGCGCCCCAAAAG 180
|||||
Db 121 CTTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGCTTTCCCGCGCCCCAAAAG 180
|||||
QY 181 AAGCGGAGATTAAAGCGGACCTGGCGGCAGAGCTGGGGAATGGCGCGCGAGCCAGGC 240
|||||
Db 181 AAGCGGAGATTAAAGCGGACCTGGCGGCAGAGCTGGGGAATGGCGCGCGAGCCAGGC 240
|||||
QY 241 CGCGCTCTCTCTCTGATGCTTTTGCAGACCGCGGTCTCGAGGGCGCTTGTGCGGT 300
|||||
Db 241 CGCGCTCTCTCTCTGATGCTTTTGCAGACCGCGGTCTCGAGGGCGCTTGTGCGGT 300
|||||
QY 301 CACACTCTCTGCACTACCTCTTTCATGGGTCCCTCAGAGCAGACCTTGGTCTTTCTTGT 360
|||||
Db 301 CACACTCTCTGCACTACCTCTTTCATGGGTCCCTCAGAGCAGACCTTGGTCTTTCTTGT 360
|||||
QY 361 TTGAGCTTTGGGCTACGTTGGATGACAGCTGTCTGTCTTCTATCATGATGAGAGTCGCC 420
|||||
Db 361 TTGAGCTTTGGGCTACGTTGGATGACAGCTGTCTGTCTTCTATCATGATGAGAGTCGCC 420
|||||

QY	421	GTCTGGAGCCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCACG	480
Db	421	GTGTGGAGCCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCACG	480
QY	481	TGAGTCAGAGTCTGAAGAGGTGGATCACATGTTCACTGTTTCTGACATTTATATGG	540
Db	481	TGAGTCAGAGTCTGAAGAGGTGGATCACATGTTCACTGTTTCTGACATTTATATGG	540
QY	541	AAATCACAACACACAGCAGAGGTCCACACCCCTCAGGTCTATCTGGGCTGTGAATGC	600
Db	541	AAATCACAACACACAGCAGAGGTCCACACCCCTCAGGTCTATCTGGGCTGTGAATGC	600
QY	601	AGAAGACACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG	660
Db	601	AGAAGACACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG	660
QY	661	AAATTCGCCCTTGACACTGGATTGGAGACAGACACCCAGGCGCTGGCCACCAAGC	720
Db	661	AAATTCGCCCTTGACACTGGATTGGAGACAGACACCCAGGCGCTGGCCACCAAGC	720
QY	721	TGGAGTGGAAAGGCACAGATTCGGGCCAGGCAGAAACAGGGCCCTACCTGGAGGGACT	780
Db	721	TGGAGTGGAAAGGCACAGATTCGGGCCAGGCAGAAACAGGGCCCTACCTGGAGGGACT	780
QY	781	GCCTGTCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
Db	781	GCCTGTCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
QY	841	CTCCCTTTGGTGAAGTGACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG	900
Db	841	CTCCCTTTGGTGAAGTGACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG	900
QY	901	CCCTGAACCTACTACCCAGACATCACCATTGAAGTGGCTGAAGATTAAGCAGCAATGG	960
Db	901	CCCTGAACCTACTACCCAGACATCACCATTGAAGTGGCTGAAGATTAAGCAGCAATGG	960
QY	961	ATGCCAAGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAAGGCT	1020
Db	961	ATGCCAAGAGTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAAGGCT	1020
QY	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACCAAGTGGAGCAC	1080
Db	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACCAAGTGGAGCAC	1080
QY	1081	CAGGCTCGATCAGCCCTCATCTGTATCTGGAGCCCTCAGCTGTGGCACCCTAGTCA	1140
Db	1081	CAGGCTCGATCAGCCCTCATCTGTATCTGGAGCCCTCAGCTGTGGCACCCTAGTCA	1140
QY	1141	TTGGAGTCATCAGTGGAAATGCTGTTTGTGCTCATCTTGTTCATTTGGAATTTGTTCA	1200
Db	1141	TTGGAGTCATCAGTGGAAATGCTGTTTGTGCTCATCTTGTTCATTTGGAATTTGTTCA	1200
QY	1201	TAATATTGAAGAGAGGAGGTTCAAGAGAGCCATGGGCACTACGTCTTAGCTGAAC	1260
Db	1201	TAATATTGAAGAGAGGAGGTTCAAGAGAGCCATGGGCACTACGTCTTAGCTGAAC	1260
QY	1261	GTGAGTCACACCGCCTGACAGTCACTGTGGGAAGAGACAAACTAGAGACTCAAG	1320
Db	1261	GTGAGTCACACCGCCTGACAGTCACTGTGGGAAGAGACAAACTAGAGACTCAAG	1320
QY	1321	AGGAGTGCAATTTATGAGCTCTTCATGTTTTCAGGAGAGGTTGAACCTAAACATAGAAAT	1380
Db	1321	AGGAGTGCAATTTATGAGCTCTTCATGTTTTCAGGAGAGGTTGAACCTAAACATAGAAAT	1380
QY	1381	TGCTGAGGAACTCTTTCATTTAGCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCCA	1440
Db	1381	TGCTGAGGAACTCTTTCATTTAGCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCCA	1440

RESULT 5
AR117795
LOCUS

AR117795 1440 bp DNA linear PAT 16-MAY-2001

DEFINITION	Sequence 11 from patent US 6140305.
ACCESSION	AR117795
VERSION	AR117795.1
KEYWORDS	GI:14098701
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 1440)
AUTHORS	Thomas, W.J., Drayna, D.T., Feder, J.N., Gnirke, A., Ruddy, D., Tsuchihashi, Z. and Wolff, R.K.
TITLE	Hereditary hemochromatosis gene products
JOURNAL	Patent: US 6140305-A 11 31-Oct-2000;
FEATURES	Location/Qualifiers 1..1440 source /organism="unknown"
BASE COUNT	347 a 354 c 408 g 331 t
ORIGIN	
Query Match 99.9%; Score 1438.4; DB 6; Length 1440;	
Best Local Similarity 99.9%; Pred. No. 0;	
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 GGGGACACTGGATCACCCTACTGTTTCAAGACGAGGTACCTTCTGCTGTAGGAGAGAGA 60
Db	1 GGGGACACTGGATCACCCTACTGTTTCAAGACGAGGTACCTTCTGCTGTAGGAGAGAGA 60
QY	61 ACTAAAGTCTGAAAGACCTGTTGCTTTCCACCAGGAAGTTTACTTGGGCATCTCCTGAG 120
Db	61 ACTAAAGTCTGAAAGACCTGTTGCTTTTCCACCAGGAAGTTTACTTGGGCATCTCCTGAG 120
QY	121 CCTAGGCAATAGCTGTAGGTGACTTCTTGAGGCCATCCCGGTTTCCCGCCCCCAAAAG 180
Db	121 CCTAGGCAATAGCTGTAGGTGACTTCTTGAGGCCATCCCGGTTTCCCGCCCCCAAAAG 180
QY	181 AAGCGAGATTTAACGGGACGTGCGGCCAGAGCTGGGGAATGGGCCCGCAGCCAGGC 240
Db	181 AAGCGAGATTTAACGGGACGTGCGGCCAGAGCTGGGGAATGGGCCCGCAGCCAGGC 240
QY	241 CGGCGCTTCTCCCTCCTGATCTTTTGACAGACCGCGTCTTGCAGAGGCGCTTCTCGCTT 300
Db	241 CGGCGCTTCTCCCTCCTGATCTTTTGACAGACCGCGTCTTGCAGAGGCGCTTCTCGCTT 300
QY	301 CACACTCTGCACTACCTCTTCATGGTCCCTCAGAGCAGGACCTTGGTCTTCTTCTTGT 360
Db	301 CACACTCTGCACTACCTCTTCATGGTCCCTCAGAGCAGGACCTTGGTCTTCTTCTTGT 360
QY	361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTCTATGATGATGAGAGTCGCC 420
Db	361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTCTGTTCTATGATGATGAGAGTCGCC 420
QY	421 GTGTGGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCACG 480
Db	421 GTGTGGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCACG 480
QY	481 TGAGTCAGAGTCTGAAGGTTGGATCAGATGTTCACTGTTTCTGACTTCTGGACTATATGG 540
Db	481 TGAGTCAGAGTCTGAAGGTTGGATCAGATGTTCACTGTTTCTGACTTCTGGACTATATGG 540
QY	541 AAAATCACAACACACAGCAAGAGTCCACACCTGCAGGTATCTTGGGCTGTGAATGC 600
Db	541 AAAATCACAACACACAGCAAGAGTCCACACCTGCAGGTATCTTGGGCTGTGAATGC 600
QY	601 AGAAGACAACTACCGAGGCTACTTGGAACTACGGGTATGATGGGAGGACCACTTGG 660
Db	601 AGAAGACAACTACCGAGGCTACTTGGAACTACGGGTATGATGGGAGGACCACTTGG 660
QY	661 AATTCTGCCCTGCACACTGGATTGGAGAGCAGCAACCCAGGCGCTACCTGGAGAGGACT 720
Db	661 AATTCTGCCCTGCACACTGGATTGGAGAGCAGCAACCCAGGCGCTACCTGGAGAGGACT 720
QY	721 TGGAGTGGAAAGGCACAGATTCGGGCCAGGCAGAAACAGGGCCCTACCTGGAGGGACT 780
Db	721 TGGAGTGGAAAGGCACAGATTCGGGCCAGGCAGAAACAGGGCCCTACCTGGAGGGACT 780

Db	721	TGAGTGGGAAAGGCACAAGATTTCGGCCAGGCAGACAGGCGCTACCTGGAGAGGGACT	780
QY	781	GCCTTGACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
Db	781	GCCTTGACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
QY	841	CTCCTTTGGTGAAGGTGACACATCATGTACCTCTTCACTGACCACTCTACGGTGTCTGGG	900
Db	841	CTCCTTTGGTGAAGGTGACACATCATGTACCTCTTCACTGACCACTCTACGGTGTCTGGG	900
QY	901	CTTTGAACCTACTACCCCGGAGACATCACCATGAAGTGGCTGAAGAGTAAGACGACCAATGG	960
Db	901	CTTTGAACCTACTACCCCGGAGACATCACCATGAAGTGGCTGAAGAGTAAGACGACCAATGG	960
QY	961	ATGCCAAGGAGTTGCAACCTTAAGAGCGTATGCCCCAATGGGATGGGACTACCGAGGCT	1020
Db	961	ATGCCAAGGAGTTGCAACCTTAAGAGCGTATGCCCCAATGGGATGGGACTACCGAGGCT	1020
QY	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATAGTGCCAGGTGGAGCACC	1080
Db	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATAGTGCCAGGTGGAGCACC	1080
QY	1081	CAGGCTGGATAGCCCTCATTTGATCTGGAGAGCCCTCAGCGTGGCACCCCTAGTCA	1140
Db	1081	CAGGCTGGATAGCCCTCATTTGATCTGGAGAGCCCTCAGCGTGGCACCCCTAGTCA	1140
QY	1141	TTGGAGTCATCAGTGGAAATGCTTTTTCGTCATCTTGTTCATTTGGAATTTTGTTC	1200
Db	1141	TTGGAGTCATCAGTGGAAATGCTTTTTCGTCATCTTGTTCATTTGGAATTTTGTTC	1200
QY	1201	TAATATTAAAGAAAGCAGCGGTTCAAGAGAGGCCATGGGCACTACGCTTTAGCTGAAC	1260
Db	1201	TAATATTAAAGAAAGCAGCGGTTCAAGAGAGGCCATGGGCACTACGCTTTAGCTGAAC	1260
QY	1261	GTAGTGACAGCGCCCTGACAGTCTACTGTGGAGGAGACAAACTAGAGACTCAAG	1320
Db	1261	GTAGTGACAGCGCCCTGACAGTCTACTGTGGAGGAGACAAACTAGAGACTCAAG	1320
QY	1321	AGGAGTGCAATTATAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT	1380
Db	1321	AGGAGTGCAATTATAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT	1380
QY	1381	TGCCTGACGAACCTCTTGAATTTAGCCTCTCTGTTCAATTTCTTCAAAAAGATTTCCCA	1440
Db	1381	TGCCTGACGAACCTCTTGAATTTAGCCTCTCTGTTCAATTTCTTCAAAAAGATTTCCCA	1440
RESULT 6			
LOCUS	AR149464	1440 bp	DNA linear PAT 08-AUG-2001
DEFINITION	Sequence 10 from patent US 6228594.		
ACCESSION	AR149464		
VERSION	AR149464.1	GI:151114055	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1440)		
AUTHORS	Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.		
TITLE	Method for determining the presence or absence of a hereditary hemochromatosis gene mutation		
JOURNAL	Patent: US 6228594-A 10 08-MAY-2001;		
FEATURES	Location/Qualifiers		
source	1..1440		
BASE COUNT	348 a	355 c	406 g
ORIGIN	331 t		
Query Match 99.98; Score 1438.4; DB 6; Length 1440;			
Best-Local Similarity 99.98; Pred. No. 0;			
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			

QY	1	GGGACACTGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTAGGAGAGAGA	60
Db	1	GGGACACTGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTAGGAGAGAGA	60
QY	61	ACTAAAGTTCTGAAGACCTGTTCTTTTACCAGGAAGTTTACTGGGATCTCCTGAG	120
Db	61	ACTAAAGTTCTGAAGACCTGTTCTTTTACCAGGAAGTTTACTGGGATCTCCTGAG	120
QY	121	CCTAGGCAATAGCTGATAGGGTACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG	180
Db	121	CCTAGGCAATAGCTGATAGGGTACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG	180
QY	181	AAGCGGAGATTAAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGCCGGGAGCCAGGC	240
Db	181	AAGCGGAGATTAAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGCCGGGAGCCAGGC	240
QY	241	CGGCGTCTCTCCTCTGATGCTTTTGCAGACCGCGGCTCTCAGGGGCGCTTGTGCGTT	300
Db	241	CGGCGTCTCTCCTCTGATGCTTTTGCAGACCGCGGCTCTCAGGGGCGCTTGTGCGTT	300
QY	301	CACACTCTCTGCACCTACCTCTTCATGGGTGCTCAGAGCAGACCTTGGTCTTCTCTGT	360
Db	301	CACACTCTCTGCACCTACCTCTTCATGGGTGCTCAGAGCAGACCTTGGTCTTCTCTGT	360
QY	361	TTGAAGCTTTTGGGCTACCTGATGACAGCTGTTGCGTGTCTATGATCATGAGAGTCGCC	420
Db	361	TTGAAGCTTTTGGGCTACCTGATGACAGCTGTTGCGTGTCTATGATCATGAGAGTCGCC	420
QY	421	GTGTGGGCGCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGGTGACG	480
Db	421	GTGTGGGCGCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGGTGACG	480
QY	481	TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGAATATGG	540
Db	481	TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGAATATGG	540
QY	541	AAATATCACACACACAGCAAGAGTCCACACCCCTGACAGTCTATGAGTGTGAAATGC	600
Db	541	AAATATCACACACACAGCAAGAGTCCACACCCCTGACAGTCTATGAGTGTGAAATGC	600
QY	601	AAGAAGACAACAGTACCAGGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG	660
Db	601	AAGAAGACAACAGTACCAGGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG	660
QY	661	AATTTCTGCCCTGACACACTTGGATTTGGAGAGCAGACCAACCCAGGCGCTGGCCCAACG	720
Db	661	AATTTCTGCCCTGACACACTTGGATTTGGAGAGCAGACCAACCCAGGCGCTGGCCCAACG	720
QY	721	TGGAGTGGGAAAGGCACACAGATTCGGGCGCAGGAGGCGCTACCTGGAGAGGACT	780
Db	721	TGGAGTGGGAAAGGCACACAGATTCGGGCGCAGGAGGCGCTACCTGGAGAGGACT	780
QY	781	GCCTTGACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
Db	781	GCCTTGACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
QY	841	CTCCTTTGGTGAAGGTGACACATCATGTACCTCTTCACTGACCACTCTACGGTGTCTGGG	900
Db	841	CTCCTTTGGTGAAGGTGACACATCATGTACCTCTTCACTGACCACTCTACGGTGTCTGGG	900
QY	901	CCTTTGAACCTACTACCCCGGAGACATCACCATGAAGTGGCTGAAGATTAAGACGCAATGG	960
Db	901	CCTTTGAACCTACTACCCCGGAGACATCACCATGAAGTGGCTGAAGATTAAGACGCAATGG	960
QY	961	ATGCCAAGGAGTTGCAACCTTAAGAGCTATTGCCCAATGGGATGGGACTACCGAGGCT	1020
Db	961	ATGCCAAGGAGTTGCAACCTTAAGAGCTATTGCCCAATGGGATGGGACTACCGAGGCT	1020
QY	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATAGTGCCAGGTGGAGCACC	1080
Db	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATAGTGCCAGGTGGAGCACC	1080

QY	1081	CAGGCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA	1140
Db	1081	CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA	1140
QY	1141	TTGGAGTCATCAGTGGAAATTCGTTTTTGTCTGTCATCTTGTTCATTTGGAAATTTGTTC	1200
Db	1141	TTGGAGTCATCAGTGGAAATTCGTTTTTGTCTGTCATCTTGTTCATTTGGAAATTTGTTC	1200
QY	1201	TAATATTAAAGAGAGCAGGGTTCAAGAGAGCCATGGGCGACTACGCTTTAGCTGAAC	1260
Db	1201	TAATATTAAAGAGAGCAGGGTTCAAGAGAGCCATGGGCGACTACGCTTTAGCTGAAC	1260
QY	1261	GTGAGTGCACAGCAGCCTGCAGACTCACTGTGGGAGGAGACAAAACACTAGAGACTCA	1320
Db	1261	GTGAGTGCACAGCAGCCTGCAGACTCACTGTGGGAGGAGACAAAACACTAGAGACTCA	1320
QY	1321	AGGAGTGCATTTATGAGCTTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAA	1380
Db	1321	AGGAGTGCATTTATGAGCTTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAA	1380
QY	1381	TGCTTGACGAACTCCTTGATTTTAGCCTTCTCTGTTTCATTTCTCAAAAAGATTTCC	1440
Db	1381	TGCTTGACGAACTCCTTGATTTTAGCCTTCTCTGTTTCATTTCTCAAAAAGATTTCC	1440
RESULT 7			
LOCUS AR149465 1440 bp DNA linear PAT 08-AUG-2001			
DEFINITION Sequence 11 from patent US 6228594.			
ACCESSION AR149465			
VERSION AR149465.1 GI:151114056			
KEYWORDS Unknown.			
SOURCE Unknown.			
ORGANISM Unclassified.			
REFERENCE 1 (bases 1 to 1440)			
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.			
TITLE Method for determining the presence or absence of a hereditary hemochromatosis gene mutation			
JOURNAL Patent: US 6228594-A 11 08-MAY-2001;			
FEATURES Location/Qualifiers			
source 1..1440			
BASE COUNT 347 a 354 c 408 g 331 t			
ORIGIN /organism="unknown"			
Query Match 99.9%; Score 1438.4; DB 6; Length 1440;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCGCTGTAGGAGAGAGA	60
Db	1	GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCGCTGTAGGAGAGAGA	60
QY	61	ACTAAAGTCTGAAGACCTTGTGCTTTTCACCAGAACTTTTACTGGGCATCTCCTGAG	120
Db	61	ACTAAAGTCTGAAGACCTTGTGCTTTTCACCAGAACTTTTACTGGGCATCTCCTGAG	120
QY	121	CCTAGGCAATAGCTGTAGGTTGACTTCTGAGAGCATCCCGCTTTCCCGCCGCCCAAA	180
Db	121	CCTAGGCAATAGCTGTAGGTTGACTTCTGAGAGCATCCCGCTTTCCCGCCGCCCAAA	180
QY	181	RAGCGGAGATTTACGGGGACGTGGCGCCAGAGCTGGGGAATGGCGCGGAGCCAGGC	240
Db	181	RAGCGGAGATTTAAGGGGACGTGGCGCCAGAGCTGGGGAATGGCGCGGAGCCAGGC	240
QY	241	CGGCGCTTCTCCTCCTGATGCTTTTGACAGACCGCGTCTCTCAGGGGCGCTTGTGCG	300
Db	241	CGGCGCTTCTCCTCCTGATGCTTTTGACAGACCGCGTCTCTCAGGGGCGCTTGTGCG	300
QY	301	CACACTCTCTGCACTACCTTCTCATGGTGCCTTCAGAGCAGGACCTTGGTCTTCTCT	360

Db	301	CACACTCTCTGCACTACCTTTCATGGTGCCTCAGAGCAGGACCTTGGTCTTCTCTTGT	360
QY	361	TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTTGGTCTTATGATCATGAGAGTCGCC	420
Db	361	TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTTGGTCTTATGATCATGAGAGTCGCC	420
QY	421	GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC	480
Db	421	GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC	480
QY	481	TCAGTCAGAGTCTGAAGGGTGGGATCACATGTTTTCACCTGTTGACTTCTGGACTATTATGG	540
Db	481	TCAGTCAGAGTCTGAAGGGTGGGATCACATGTTTTCACCTGTTGACTTCTGGACTATTATGG	540
QY	541	AAATTCACAAACACACAGCAGGAGTCCCACACCTCCAGGTTCATCCTGGGCTGTGAATGC	600
Db	541	AAATTCACAAACACACAGCAGGAGTCCCACACCTCCAGGTTCATCCTGGGCTGTGAATGC	600
QY	601	AAGAAGCAACAGTACCGAGGGCTACTGGAGTACGGGTATGATGGCGAGGACCACTTG	660
Db	601	AAGAAGCAACAGTACCGAGGGCTACTGGAGTACGGGTATGATGGCGAGGACCACTTG	660
QY	661	AATTCTGCCCTGACACACTGGATTGGAGCAGCAGAACCCAGGCGCTGGCCCCAACAGC	720
Db	661	AATTCTGCCCTGACACACTGGATTGGAGCAGCAGAACCCAGGCGCTGGCCCCAACAGC	720
QY	721	TGGAGTGGGAAAGGCACAAGATTTCGGGCCAGGCACAACAGGCGCTTACCTGGAGAGGACT	780
Db	721	TGGAGTGGGAAAGGCACAAGATTTCGGGCCAGGCACAACAGGCGCTTACCTGGAGAGGACT	780
QY	781	GCCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
Db	781	GCCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
QY	841	CTCCTTTTGGTGAAGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGCGG	900
Db	841	CTCCTTTTGGTGAAGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGCGG	900
QY	901	CCTTGAACCTACTACCCCCAGAACATCACCATGAAGTGTCTGAAGGATAAGCAGCAATGG	960
Db	901	CCTTGAACCTACTACCCCCAGAACATCACCATGAAGTGTCTGAAGGATAAGCAGCAATGG	960
QY	961	ATGCCAAGAGGTTTCAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT	1020
Db	961	ATGCCAAGAGGTTTCAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT	1020
QY	1021	GGATAACCTTTGGCTGTACCCCTGGGGAAGACGAGATATACGTGCCAGCTGGAGCACC	1080
Db	1021	GGATAACCTTTGGCTGTACCCCTGGGGAAGACGAGATATACGTGCCAGCTGGAGCACC	1080
QY	1081	CAGGCTTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA	1140
Db	1081	CAGGCTTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA	1140
QY	1141	TTGAGTCAATCAGTGAATTCGCTGTTTTCGTCATCTTGTTCATTTGGAAATTTGTTCA	1200
Db	1141	TTGAGTCAATCAGTGAATTCGCTGTTTTCGTCATCTTGTTCATTTGGAAATTTGTTCA	1200
QY	1201	TAATATTAAAGAGAGCAGGTTTCAAGAGAGCCATGGGCGACTACGCTTTAGCTGAAC	1260
Db	1201	TAATATTAAAGAGAGCAGGTTTCAAGAGAGCCATGGGCGACTACGCTTTAGCTGAAC	1260
QY	1261	GTGAGTCAACAGCAGCCTTGCAGACTCACTGTGGGAAGSAGACAAAACCTAGAGACTCAAAG	1320
Db	1261	GTGAGTCAACAGCAGCCTTGCAGACTCACTGTGGGAAGSAGACAAAACCTAGAGACTCAAAG	1320
QY	1321	AGGAGTGCATTTATGAGCTTCTCATGTTTTCAGGAGAGAGTTGAACCTTAACATAGAAAT	1380
Db	1321	AGGAGTGCATTTATGAGCTTCTCATGTTTTCAGGAGAGAGTTGAACCTTAACATAGAAAT	1380
QY	1381	TGCTGACGAACTCCTTGTGATTTTAGCCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCCA	1440
Db	1381	TGCTGACGAACTCCTTGTGATTTTAGCCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCCA	1440

Db 1381 TGCCCTGACGAACTCCTTGATTTTAGCCTTCTCTGTTCATTTCTCTCAAAAAGATTTCCCA 1440

RESULT 8

LOCUS AR117796 1440 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 12 from patent US 6140305.

ACCESSION AR117796

VERSION AR117796.1 GI:14098702

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1440)

AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.

TITLE Hereditary hemochromatosis gene products

JOURNAL Patent: US 6140305-A 12 31-Oct-2000;

FEATURES Location/Qualifiers

1..1440

/organism="unknown"

BASE COUNT 348 a 354 c 407 g 331 t

ORIGIN

Query Match 99.8%; Score 1436.8; DB 6; Length 1440;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGTACCTTCTGTGTAGGAGAGAGA 60

Db 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGTACCTTCTGTGTAGGAGAGAGA 60

Qy 61 ACTAAAGTCTCTGAAGACCTTGTGCTTTTACCAGGAAGTTTACGTGGCATCTCTGAG 120

Db 61 ACTAAAGTCTCTGAAGACCTTGTGCTTTTACCAGGAAGTTTACGTGGCATCTCTGAG 120

Qy 121 CCTAGCAATAGCTGTAGGTGACTTCTGTAGCCATCCCGTTCGCCGCCGCCGAGAG 180

Db 121 CCTAGCAATAGCTGTAGGTGACTTCTGTAGCCATCCCGTTCGCCGCCGCCGAGAG 180

Qy 181 AAGCGGAGATTTAACGGGACGTGCGGCCAGAGCTGGGGAATGGCCGCCGAGCAGGC 240

Db 181 AAGCGGAGATTTAACGGGACGTGCGGCCAGAGCTGGGGAATGGCCGCCGAGCAGGC 240

Qy 241 CGGCGCTTCTCCCTGATGCTTTTGACAGCCGCTGCTGACGGGCGCTTGCTCGTT 300

Db 241 CGGCGCTTCTCCCTGATGCTTTTGACAGCCGCTGCTGACGGGCGCTTGCTCGTT 300

Qy 301 CACACTCTGCACTACCTTCTCATGGGTGCTGACAGCAGGACCTTGTGCTTCTGTT 360

Db 301 CACACTCTGCACTACCTTCTCATGGGTGCTGACAGCAGGACCTTGTGCTTCTGTT 360

Qy 361 TTGAAGCTTTGGGCTAGTGATGACAGCTGTCGTGTCTATGATCATGAGATCGCC 420

Db 361 TTGAAGCTTTGGGCTAGTGATGACAGCTGTCGTGTCTATGATGATGAGATCGCC 420

Qy 421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Db 421 GTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTGTGACTTCTGACTATTATGG 540

Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTGTGACTTCTGACTATTATGG 540

Qy 541 AAAATCACACACAGCAAGAGTCCACACCTGAGTCATCTCCGGCTGTGAATGC 600

Db 541 AAAATCACACACAGCAAGAGTCCACACCTGAGTCATCTCCGGCTGTGAATGC 600

Qy 601 AAGAAGACAACAGTACCGAGGCGTACTTGGGAAGTACGGGTATGATGGCAGGACCACTTG 660

Db 601 AAGAAGACAACAGTACCGAGGCGTACTTGGGAAGTACGGGTATGATGGCAGGACCACTTG 660

Qy 661 AATTCTGCCCTGACACACTGGATTTGGAGAGCAGCAGAACCCAGGGCCTGGCCCAACAGC 720

Db 661 AATTCTGCCCTGACACACTGGATTTGGAGAGCAGCAGAACCCAGGGCCTGGCCCAACAGC 720

Qy 721 TGGAGTGGGAAGGCACAAAGATTCGGGCCAGGCAGACAGGGCCTACTCTGGAGAGGACT 780

Db 721 TGGAGTGGGAAGGCACAAAGATTCGGGCCAGGCAGACAGGGCCTACTCTGGAGAGGACT 780

Qy 781 GCCCTGACAGCTGCAGCAGTTCGTGGAGCTGGGGAGAGTGTTTTGGACCAACAATGC 840

Db 781 GCCCTGACAGCTGCAGCAGTTCGTGGAGCTGGGGAGAGTGTTTTGGACCAACAATGC 840

Qy 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTTACGGTGTGGG 900

Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTTACGGTGTGGG 900

Qy 901 CTTTGAACCTACTACCCCGAGAACATCACATGAAGTGGCTGAAGGATAAGCAGCAATGG 960

Db 901 CTTTGAACCTACTACCCCGAGAACATCACATGAAGTGGCTGAAGGATAAGCAGCAATGG 960

Qy 961 ATGCCAAGGAGTTGGAACCTAAAGACGTATTGGCCCAATGGGGATGGACCTTACCAAGGCT 1020

Db 961 ATGCCAAGGAGTTGGAACCTAAAGACGTATTGGCCCAATGGGGATGGACCTTACCAAGGCT 1020

Qy 1021 GGATTAACCTTTGGCTGTACCCCTGGGGAAGAGAGAGATATACGTGCCAGTGAGCACC 1080

Db 1021 GGATTAACCTTTGGCTGTACCCCTGGGGAAGAGAGAGATATACGTGCCAGTGAGCACC 1080

Qy 1081 CAGGCTCGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Db 1081 CAGGCTCGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy 1141 TTGAGTGCATCAGTGGAAATTCGTGTTTGTGTCATCTTGTTCATTGGAAATTTGTTC 1200

Db 1141 TTGAGTGCATCAGTGGAAATTCGTGTTTGTGTCATCTTGTTCATTGGAAATTTGTTC 1200

Qy 1201 TAATATTAAAGAAAGGAGGAGGTTCAAGAGAGCCTAGGGCAGTCTTACCTGAC 1260

Db 1201 TAATATTAAAGAAAGGAGGAGGTTCAAGAGAGCCTAGGGCAGTCTTACCTGAC 1260

Qy 1261 GTGAGTGACAGCAGCCTGCAGACTCACTGTGGGAAGAGACAAAACCTAGAGACTCAAG 1320

Db 1261 GTGAGTGACAGCAGCCTGCAGACTCACTGTGGGAAGAGACAAAACCTAGAGACTCAAG 1320

Qy 1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTTAAACATAGAAAT 1380

Db 1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTTAAACATAGAAAT 1380

Qy 1381 TGCCCTGACGAACCTCCTTGAATTTAGCCTTCTGTTTCATTTCCCTCAAAAAGATTTCCCA 1440

Db 1381 TGCCCTGACGAACCTCCTTGAATTTAGCCTTCTGTTTCATTTCCCTCAAAAAGATTTCCCA 1440

RESULT 9

AR149466

LOCUS AR149466 1440 bp DNA linear PAT 08-AUG-2001

DEFINITION Sequence 12 from patent US 6228594.

ACCESSION AR149466

VERSION AR149466.1 GI:15114057

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1440)

AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.

TITLE Method for determining the presence or absence of a hereditary hemochromatosis gene mutation

JOURNAL Patent: US 6228594-A 12 08-MAY-2001;

FEATURES Location/Qualifiers

1..1440

/organism="unknown"

BASE COUNT 348 a 354 c 407 g 331 t

ORIGIN

Query Match 99.8%; Score 1436.8; DB 6; Length 1440;
Best Local Similarity 99.9%; Mismatches 0; Indels 0; Gaps 0;
Matches 1438; Conservative 0; P1: 0;

QY 1 GGGGACATGGATACCTAGTGTTCACAAAGCAGGTACCTCTGCTGTAGGAGAGAGA 60
DB 1 GGGGACATGGATACCTAGTGTTCACAAAGCAGGTACCTCTGCTGTAGGAGAGAGA 60
QY 61 ACTAAAGTCTGAAGACCTGTGCTTTTACACAGAGTCTTACTTGGGCACTCTCGTGA 120
DB 61 ACTAAAGTCTGAAGACCTGTGCTTTTACACAGAGTCTTACTTGGGCACTCTCGTGA 120
QY 121 CTTAGGCAATAGCTGTAGGAGTGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAG 180
DB 121 CTTAGGCAATAGCTGTAGGAGTGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAG 180
QY 181 AAGCGGAGATTTAAGCGGACGTGCGGCGAGAGCTGGGGAATGGGCGCGAGCGAGGC 240
DB 181 AAGCGGAGATTTAAGCGGACGTGCGGCGAGAGCTGGGGAATGGGCGCGAGCGAGGC 240
QY 241 CGGCGCTTCTCTCTGATGCTTTTTCAGACCGCGTCTGAGGAGGCGCTTCTGCGGTT 300
DB 241 CGGCGCTTCTCTCTGATGCTTTTTCAGACCGCGTCTGAGGAGGCGCTTCTGCGGTT 300
QY 301 CACACTCTCTGCACTACCTCTTTCATGGGTCCTTCAGAGCAGGACCTTGGTCTTCTTGT 360
DB 301 CACACTCTCTGCACTACCTCTTTCATGGGTCCTTCAGAGCAGGACCTTGGTCTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTTCTGTTCTATGATCATGAGTCGCC 420
DB 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTTCTGTTCTATGATCATGAGTCGCC 420
QY 421 GTGTGGAGCCCCCACTCCAGTGGTTCAGATAGATTTCAAGCCAGATGTGCTGCAGC 480
DB 421 GTGTGGAGCCCCCACTCCAGTGGTTCAGATAGATTTCAAGCCAGATGTGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGTTGGGATCACATGTTTCACTGTTGACTTCTGACTATTATGG 540
DB 481 TGAGTCAGAGTCTGAAGGTTGGGATCACATGTTTCACTGTTGACTTCTGACTATTATGG 540
QY 541 AAAATCACACACAGGAGTCCACACCTGAGTATGATGGGCTGTGAATGC 600
DB 541 AAAATCACACACAGGAGTCCACACCTGAGTATGATGGGCTGTGAATGC 600
QY 601 AAGAGACACAGTACCGAGGCTACTTGGAGTACGGGTATGATGGGCGAGGACCTTGG 660
DB 601 AAGAGACACAGTACCGAGGCTACTTGGAGTACGGGTATGATGGGCGAGGACCTTGG 660
QY 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGACACCCAGGGCTTGGCCCAACAGC 720
DB 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGACACCCAGGGCTTGGCCCAACAGC 720
QY 721 TGGAGTGGGAAAGGCACAAATTCGGGCGAGGAGGAGGCTACCTGGAGGAGGACT 780
DB 721 TGGAGTGGGAAAGGCACAAATTCGGGCGAGGAGGAGGCTACCTGGAGGAGGACT 780
QY 781 GCCCTGCACAGCTGCAGCAGTGTGCTGGAGCTGGGAGAGGCTGTTTGGACCAACAAGTGC 840
DB 781 GCCCTGCACAGCTGCAGCAGTGTGCTGGAGCTGGGAGAGGCTGTTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTACCTTCTTCACTGACCACTTCTACGGTGTGGG 900
DB 841 CTCCTTTGGTGAAGGTGACACATCATGTACCTTCTTCACTGACCACTTCTACGGTGTGGG 900
QY 901 CTTTGAACCTACTACCCCGAACAATCACCATGAGTGGCTGAAGGATTAAGCAGGCAATGG 960
DB 901 CTTTGAACCTACTACCCCGAACAATCACCATGAGTGGCTGAAGGATTAAGCAGGCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTTAAGACCTATTGCCAATGGGATGGGACCTACCAAGGCT 1020
DB 961 ATGCCAAGGAGTTCGAACCTTAAGACCTATTGCCAATGGGATGGGACCTACCAAGGCT 1020

Db 961 ATGCCAAGGAGTTCGAACCTTAAGACCTATTGCCAATGGGATGGACCTACCAAGGCT 1020
QY 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGAGCAGATATAGTGTCCAGGTGGAGCAC 1080
Db 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGAGCAGATATAGTGTCCAGGTGGAGCAC 1080
QY 1081 CAGCCCTGGATCAGCCCTCATTTGTATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db 1081 CAGCCCTGGATCAGCCCTCATTTGTATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
QY 1141 TTGGAGTCATCAGTGAATTCCTTTTGTGTCATCTTGTTCATTTGGAATTTGTTC 1200
Db 1141 TTGGAGTCATCAGTGAATTCCTTTTGTGTCATCTTGTTCATTTGGAATTTGTTC 1200
QY 1201 TAATATTAAAGGAAGAGCAGGTTCAAGAGAGAGCCATGGGCGACTAGCTCTTACCTGAAC 1260
Db 1201 TAATATTAAAGGAAGAGCAGGTTCAAGAGAGAGCCATGGGCGACTAGCTCTTACCTGAAC 1260
QY 1261 GTGAGTCACAGCCCTGCAGACTCACTGTGGGAAGAGACAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTCACAGCCCTGCAGACTCACTGTGGGAAGAGACAAACTAGAGACTCAAG 1320
QY 1321 AGGAGTGCATTTATGAGCTCTTCACTGTTTTCAGGAGAGTGAACCTTAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTCTTCACTGTTTTCAGGAGAGTGAACCTTAACATAGAAAT 1380
QY 1381 TGCCTGACGAACCTCTTGAATTTAGCCTTCTCTTTCATTTCTCCCAAAAGATTCCCA 1440
Db 1381 TGCCTGACGAACCTCTTGAATTTAGCCTTCTCTTTCATTTCTCCCAAAAGATTCCCA 1440
RESULT 10
HSA249335 1280 bp mRNA linear PRI 04-SEP-2001
LOCUS Homo sapiens mRNA for hemochromatosis protein (HFE gene) splice
DEFINITION variant 1.
ACCESSION AJ249335
VERSION AJ249335.1 GI:15485418
KEYWORDS alternative splicing; hemochromatosis protein; HFE gene.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1280)
AUTHORS Oliva R. and Sanchez M.
TITLE Identification of different alternative splicing forms of the HFE gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1280)
AUTHORS Oliva R.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-1999) Oliva R., Faculty of Medicine and Clinic Hospital, Human Genome Research Group, Casanova 143, 08036, SPAIN
FEATURES
1..1280 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="p22"
/cell_line="HepG2"
161..1138
/gene="HFE"
161..1138
/gene="HFE"
/function="iron metabolism"
/note="alternative splicing form with deletion of first 69 pb of exon 2"
/codon_start=1
/evidence="experimental"
/product="hemochromatosis protein"
/protein_id="CAC67792.1"
/db_xref="GI:15485419"
/translation="MGPRARPALLLLMLLTAVLQGRLLPLGYVDDQLFVYDHSRR"

VEPRTPVSSRISSOMWLQSLKQWDMHFTVDFWTIMENHNHRSKESHTLOVILGCE
MOENSTEGYWKYCGDQHLFECPDPTLDWRAAEPRAWPTKLEWRHKIRARONRAVL
ERDCPAOQLLELGRVLDOQVPLVKVTHVTSVVTLLRCALANYPONTIMKWLK
DKOPMDAKEPEKDVLPNGDGTYGWHTITLAVPGEBQRTYTCQVEHFGDQPLIVIEWP
SPSGTLVIGVISGIAVNFVILFGLILFILLRKQSGRGAMGHVYLAERE"
BASE COUNT 311 a 314 c 371 g 284 t
ORIGIN

Query Match 83.48; Score 1201; DB 9; Length 1280;
Best Local Similarity 94.99; Pred. No. 0;
Matches 1280; Conservative 0; Mismatches 0; Indels 69; Caps 1;
Qy 62 CTAAGTTCTGAAAGACCTGTGCTTTTCCACGAGAAATTTACTGGGCATCTCCCTGAGC 121
Db 1 CTAAGTTCTGAAAGACCTGTGCTTTTCCACGAGAAATTTACTGGGCATCTCCCTGAGC 60
Qy 122 CTAGGCAATAGCTGTAGGGTACTTCTGGAGCCATCCCGCTTTCCCGCCCCCAAAAGA 181
Db 61 CTAGGCAATAGCTGTAGGGTACTTCTGGAGCCATCCCGCTTTCCCGCCCCCAAAAGA 120
Qy 182 AGCGAGATTTAACGGGAGCTGCGGCCAGAGCTGGGAAATGGGCCCGCGAGCCAGGCC 241
Db 121 AGCGAGATTTAACGGGAGCTGCGGCCAGAGCTGGGAAATGGGCCCGCGAGCCAGGCC 180
Qy 242 GGCCTTCTCCTCTGATGCTTTTGCAGACCGGGCTCCTGCAGGGCGCTTGCCTGCGTTC 301
Db 181 GGCCTTCTCCTCTGATGCTTTTGCAGACCGGGCTCCTGCAGGGCGCTTGCCTGCGTTC 236
Qy 302 ACATCTCTGCACACTACCTCTTCATGGTGGCTCAGAGCAGGACCTTGGCTTTTCCTTGT 361
Db 237 ----- 236
Qy 362 TGAAGCTTTGGGCTACGTGGATGACGAGCTGTTCTGTTCTATGATCATGAGAGTCGCCG 421
Db 237 -----CTTTGGGCTACGTGGATGACGAGCTGTTCTGTTCTATGATCATGAGAGTCGCCG 291
Qy 422 TGTGGAGCCCGAACTCCATGGGTTTCCAGTAAATTTCAAGCCAGATGTGGCTGCAGCT 481
Db 292 TGTGGAGCCCGAACTCCATGGGTTTCCAGTAAATTTCAAGCCAGATGTGGCTGCAGCT 351
Qy 482 GAGTCAGAGTCTGAAGGGTGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGCA 541
Db 352 GAGTCAGAGTCTGAAGGGTGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGCA 411
Qy 542 AAATCAACACCAAGAGAGTCCACACCCCTGCAGGTCATCTCGGGCTGTGAATGCA 601
Db 412 AAATCAACACCAAGAGAGTCCACACCCCTGCAGGTCATCTCGGGCTGTGAATGCA 471
Qy 602 AGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCCAGGACCACTTGA 661
Db 472 AGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCCAGGACCACTTGA 531
Qy 662 ATTCTGCCCTGACACACTGATTTGGAGAGCAGAGCAACCCAGGCGCTGGCCCAACCAAGCT 721
Db 532 ATTCTGCCCTGACACACTGATTTGGAGAGCAGAGCAACCCAGGCGCTGGCCCAACCAAGCT 591
Qy 722 GGAGTGGGAAGCCACAAGATTCGGGCCAGGAGGACACAGGCGCTACCTGGAGAGGACTG 781
Db 592 GGAGTGGGAAGCCACAAGATTCGGGCCAGGAGGACACAGGCGCTACCTGGAGAGGACTG 651
Qy 782 CCCTGCACAGCTGCAGAGTTCCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 841
Db 652 CCCTGCACAGCTGCAGAGTTCCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 711
Qy 842 TCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTTACGGTGTGGGC 901
Db 712 TCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTTACGGTGTGGGC 771
Qy 902 CTTGAAGTACTACCCCAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGA 961
Db 772 CTTGAAGTACTACCCCAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGA 831

Qy 962 TGCCAAGGAGTTCGAACCTAAAGACGCTATTGCCCATGGGATGGACCTACGAGGGCTG 1021
Db 832 TGCCAAGGAGTTCGAACCTAAAGACGCTATTGCCCATGGGATGGACCTACGAGGGCTG 891
Qy 1022 GATAACCTTTGGCTGTACCCCTTGGGGAAGAGCAGATATACGTGCCAGGTGGAGCACCC 1081
Db 892 GATAACCTTTGGCTGTACCCCTTGGGGAAGAGCAGATATACGTGCCAGGTGGAGCACCC 951
Qy 1082 AGGCTGGATCAGCCCTCATTTGTGATCTGGAGCCCTCACCTCTGGCACCTAGTCAT 1141
Db 952 AGGCTGGATCAGCCCTCATTTGTGATCTGGAGCCCTCACCTCTGGCACCTAGTCAT 1011
Qy 1142 TGGAGTCATCAGTGGAAATGCTTTTGTCTCATCTTGTTCATTTGGAATTTTGTTCAT 1201
Db 1012 TGGAGTCATCAGTGGAAATGCTTTTGTCTCATCTTGTTCATTTGGAATTTTGTTCAT 1071
Qy 1202 AATATTAAGGAAGAGCAGGTTTCAAGAGAGCCACTGGGGCACTACGCTTACGCTGAACG 1261
Db 1072 AATATTAAGGAAGAGCAGGTTTCAAGAGAGCCACTGGGGCACTACGCTTACGCTGAACG 1131
Qy 1262 TGAGTGACACGAGCAGCTGCAGACTCACCTGTGGGAAGAGACAAAACCTAGAGACTCAAGA 1321
Db 1132 TGAGTGACACGAGCAGCTGCAGACTCACCTGTGGGAAGAGACAAAACCTAGAGACTCAAGA 1191
Qy 1322 GGGAGTGCAATTTATGAGCTCTTCATGTTTTCAGGAGAGATTCAACCTAAACATAGAAAT 1381
Db 1192 GGGAGTGCAATTTATGAGCTCTTCATGTTTTCAGGAGAGATTCAACCTAAACATAGAAAT 1251
Qy 1382 GCCTGACGAACCTCTTGATTTTAGCCTTC 1410
Db 1252 GCCTGACGAACCTCTTGATTTTAGCCTTC 1280

RESULT 11
AF115265
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
CDS
1200 bp mRNA linear PRI 07-MAY-2001
Homo sapiens hemochromatosis termination variant tereE (HFE) mRNA,
complete cds.
AF115265
AF115265.1 GI:11094314
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Thenie,A., Orhant,M., Gicquel,I., Fergelot,P., Le Gall,J.Y.,
David,V. and Mosser,J.
The HFE gene undergoes alternate splicing processes
Blood Cells Mol. Dis. 26 (2), 155-162 (2000)
20448010
11001625
2 (bases 1 to 1200)
Thenie,A., Orhant,M. and Mosser,J.
Direct Submission
Submitted (17-DEC-1998) UPR 41 CNRS, Faculte de Medecine, 2, av du
Pr. Bernard, Rennes 35043, France
Location/Qualifiers
1..1200
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p22"
1..1200
/gene="HFE"
12..1058
/gene="HFE"
/codon_start=1
/product="hemochromatosis termination variant tereE"
/protein_id="AAG29572.1"
/db_xref="GI:11094315"
/translation="MGPRAPALLLLMLLQTAVLQGRLLRSLHSLYLFPMGASEFDLGL
SLEFALGYDDQLFVFDHESRVEPTPWSSRISSQMWLQLSQSLKQWDMHFTVDF

WTIMENHNHSHKESHFTLOVILGCEMQEDNSTEGYWKYGDQDHLERCPDITLDHRAEP
RAWPTKLEWRHRKIRARONRAYLERDCPAQLQLELGRGVLDQVPLVYKVTYHHRVIS
SVYTLRICALINYIPONTIMKWKDKPMDAKEFEPKDVLPNGDGTGOGWITLAVPPGE
EQRYTCQVHEFLGLDQPLIVIEWPSPSGTLVIGVISGIAVFWVILFIGILFILLRKRQG
SRGANGHYVLAERE"

BASE COUNT 298 a 290 c 346 g 266 t
ORIGIN

Query Match 82.6%; Score 1189; DB 9; Length 1200;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 GAGCTGGGAAATGGCCCGCAGAGCAGGCCGCGCTCTCCTCGATGCTTTTGACA 270
DB 1 GAGCTGGGAAATGGCCCGCAGAGCAGGCCGCGCTCTCCTCGATGCTTTTGACA 60
QY 271 CCGCGTCTCGAGGGCGCTTGTGCGTTTACACTCTCTGACCTACCTCTTCATGGTG 330
DB 61 CCGCGTCTCGAGGGCGCTTGTGCGTTTACACTCTCTGACCTACCTCTTCATGGTG 120
QY 331 CCTCAGCAGGACCTTGGCTTTCTTCTTTTGAAGCTTTGGCTACGTCGATGACCAAGC 390
DB 121 CCTCAGCAGGACCTTGGCTTTCTTCTTTTGAAGCTTTGGCTACGTCGATGACCAAGC 180
QY 391 TGTTCGTGTTATGATCATGAGAGTCGCGGTGAGGCCCGGAACCTCCATGGGTTTCCA 450
DB 181 TGTTCGTGTTATGATCATGAGAGTCGCGGTGAGGCCCGGAACCTCCATGGGTTTCCA 240
QY 451 GTAGATTTCAAGCCAGATGTGCTCAGCTGAGTCAGAGTCGTAAGGGTGGATCACA 510
DB 241 GTAGATTTCAAGCCAGATGTGCTCAGCTGAGTCAGAGTCGTAAGGGTGGATCACA 300
QY 511 TGTTCAGTGTGACTTCTTGACTATTATGAAATATCACAACACAGCAGGATCCCA 570
DB 301 TGTTCAGTGTGACTTCTTGACTATTATGAAATATCACAACACAGCAGGATCCCA 360
QY 571 CCTTCAGGTCATCTGGGCTGTAATCAAGACACACAGTACCGAGGCTACTGGA 630
DB 361 CCTTCAGGTCATCTGGGCTGTAATCAAGACACACAGTACCGAGGCTACTGGA 420
QY 631 AGTACGGGTATGATGGCAGGACCACTTGAATTTCTGCCCTGACACACTGGATTGGAG 690
DB 421 AGTACGGGTATGATGGCAGGACCACTTGAATTTCTGCCCTGACACACTGGATTGGAG 480
QY 691 CAGCAGAACCCAGGCTTGGCCCAACAGCTGGAGTGGAAAGCACAAGATTGGGGCA 750
DB 481 CAGCAGAACCCAGGCTTGGCCCAACAGCTGGAGTGGAAAGCACAAGATTGGGGCA 540
QY 751 GGCAGAACAGGGCTACCTGGAGAGGACTGGCCCTGCACAGCTGCAGAGTTGCTGGAGC 810
DB 541 GGCAGAACAGGGCTACCTGGAGAGGACTGGCCCTGCACAGCTGCAGAGTTGCTGGAGC 600
QY 811 TGGGAGAGGTGTTTGGACCAACAGTGCCTCTTGTGTAAGTGACACATCATGTGA 870
DB 601 TGGGAGAGGTGTTTGGACCAACAGTGCCTCTTGTGTAAGTGACACATCATGTGA 660
QY 871 CCTCTTCAGTGACACTCTACGCTGTGCGGCTTGAACCTACCTACCCAGAACATCACA 930
DB 661 CCTCTTCAGTGACACTCTACGCTGTGCGGCTTGAACCTACCTACCCAGAACATCACA 720
QY 931 TGAAGTGGCTGAAGGATAAGCAGCCCAATGGATGCCAAGGAGTTGCAACCTAAAGACGTAT 990
DB 721 TGAAGTGGCTGAAGGATAAGCAGCCCAATGGATGCCAAGGAGTTGCAACCTAAAGACGTAT 780
QY 991 TGCCCAATGGGATGGGACTACCGGCTGGATACCTTGGCTTACCCCTGGGGAAG 1050
DB 781 TGCCCAATGGGATGGGACTACCGGCTGGATACCTTGGCTTACCCCTGGGGAAG 840
QY 1051 AGCAGAGATATAGTGTCCAGGTGGAGCACCAGGCTGGATGAGCCCTCATTTGTGATCT 1110
DB 841 AGCAGAGATATAGTGTCCAGGTGGAGCACCAGGCTGGATGAGCCCTCATTTGTGATCT 900

QY 1111 GGGAGCCCTCACCGTCTGCCACCCCTAGTTCATGGAGTTCATCAGTGAATTCCTGTTTTG 1170
DB 901 GGGAGCCCTCACCGTCTGCCACCCCTAGTTCATGGAGTTCATCAGTGAATTCCTGTTTTG 960
QY 1171 TCGTCATCTTGTTCATTCGAAATTTTGTTCATAATATTAAAGGAAGCAGGCTTCAAGAG 1230
DB 961 TCGTCATCTTGTTCATTCGAAATTTTGTTCATAATATTAAAGGAAGCAGGCTTCAAGAG 1020
QY 1231 GAGCCATGGGCACCTACCTCTTAGCTGAAGTGTGAGTGCACGCGCTGCAGACTCACTG 1290
DB 1021 GAGCCATGGGCACCTACCTCTTAGCTGAAGTGTGAGTGCACGCGCTGCAGACTCACTG 1080
QY 1291 TGGGAAGGAGACAAACTAGAGACTCAAGAGGGAGTGCATTTATGAGCTCTTCATGTTT 1350
DB 1081 TGGGAAGGAGACAAACTAGAGACTCAAGAGGGAGTGCATTTATGAGCTCTTCATGTTT 1140
QY 1351 CAGGAGAGGTGGAACCTAAACATAGAAATTCCTGAGCAACTCCTTGA 1399
DB 1141 CAGGAGAGGTGGAACCTAAACATAGAAATTCCTGAGCAACTCCTTGA 1189
RESULT 12
AF079407 1045 bp mRNA linear PRI 18-MAR-1999
LOCUS Homo sapiens hemochromatosis splice variant dell4E4 (HFE) mRNA,
DEFINITION complete cds.
ACCESSION AF079407
VERSION AF079407.1 GI:3695106
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS Rhodes,D.A. and Trowsdale,J.
TITLE Alternate splice variants of the hemochromatosis gene Hfe
JOURNAL Immunogenetics 49 (4), 357-359 (1999)
MEDLINE 99180629
REFERENCE 2 (bases 1 to 1045)
AUTHORS Rhodes,D.A.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1998) Immunology, University of Cambridge, Tennis
Court Road, Cambridge CB2 1QP, UK
FEATURES
Location/Qualifiers
1..1045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p22.1"
1..1045
/gene="HFE"
37..1041
/gene="HFE"
/cdoon_start=1
/product="hemochromatosis splice variant dell4E4"
/protein_id="AAC62646.1"
/db_xref="GI:3695107"
/translation="MGRARPALILLMLLOTAVLQGLRLSHSLHYLFMGASPDQLGL
SLFEALGYDDQLFVYDHSRRVEPTPWSSRISSOMLQLSLSKGMWHMFTVDF
WTIMENHNHSHKESHFTLOVILGCEMQEDNSTEGYWKYGDQDHLERCPDITLDHRAEP
RAWPTKLEWRHRKIRARONRAYLERDCPAQLQLELGRGVLDQVPLVYKVTYHHRANLYP
ONITMKLWKDKPMDAKEFEPKDVLPNGDGTGOGWITLAVPPGEQRYTCQVHEFLGLD
QPLVIEWPSPSGTLVIGVISGIAVFWVILFIGILFILLRKRGRSGANGHYVLAERE"

BASE COUNT 243 a 259 c 314 g 229 t
ORIGIN

Query Match 69.0%; Score 993; DB 9; Length 1045;
Best Local Similarity 96.1%; Pred. No. 2.7e-280;
Matches 1045; Conservative 0; Mismatches 0; Indels 42; Gaps 1;

QY 186 GAGATTACGGGGACGTGGCCGACGAGCTGGGAAATGGGCCCGCAGCCAGCCGCGC 245

Db	1039	TCACAGC	1045	
RESULT	13			
LOCUS	AY007541	1320 bp	mRNA	linear
DEFINITION	Ceratothorium simum HFE protein mRNA, complete cds.			
ACCESSION	AY007541			
VERSION	AY007541.1	GI:10945687		
KEYWORDS	white rhinoceros.			
SOURCE	ORGANISM Ceratothorium simum			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.			
REFERENCE	1	(bases 1 to 1320)		
AUTHORS	West,C.J., Worley,M. and Beutler,E.			
TITLE	Rhinoceros HFE Polymorphisms			
JOURNAL	Unpublished			
REFERENCE	2	(bases 1 to 1320)		
AUTHORS	West,C.J., Worley,M. and Beutler,E.			
TITLE	Direct Submission			
JOURNAL	Submitted (29-AUG-2000) Molecular and Experimental Medicine, The			
	Scripps Research Institute, 10550 North Torrey Pines Road, La			
	Jolla, CA 92037, USA			
FEATURES	Location/Qualifiers			
source	1..1320	/organism="Ceratothorium simum"		
		/db_xref="taxon:9807"		
CDS	86..1132	/codon_start=1		
		/product="HFE protein"		
		/protein_id="AAG23701.1"		
		/db_xref="GI:10945688"		
		/translation="MGPRPALFFLLRLTVAAQRRPRSHSLRYLFMGASERDHLG		
		PLFALGVYDDELVAYNHESRRAESRAQWLGHAQSLQWLQSQSLKGDHDFIVDF		
		WTIMDNHNSKSHLTQVILGCEVEDNSTGFWKYGVDGQDHLFEFCPETLDMRAAES		
		RALTTLKEWYNKIRAKONRAYLERDCPEQLQWLLLELGRVLDQVPLVKVTHHVAS		
		AVTTLRQALNFYQNTITMRWLKDKRPMVDKDAESKDVLPDGDGTYSQSWAALAVPPE		
		EQNYTCQVEHPGLDPLTATWELPSLSLNTLVGTVISGIAVCVLIFFIGILFRILKKROA		
		SRGAMGDVILGECE"		
BASE COUNT	284 a	363 c	394 g	279 t
ORIGIN				
Query Match	62.3%;	Score 897;	DB 4;	Length 1320;
Best Local Similarity	82.4%;	Pred. No. 4.5e-252;		
Matches 1065;	Conservative	0;	Mismatches 225;	Indels 3; Gaps 3;
QY	137	AGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCGCCCCCAAAAGAGCGGAGATTAAAG	196	
Db	2	AGAGTGACTTCTGGGTCACTCCTATTTCGCCGCCCCACCGAAGCAGCTGA-CGGCGGCC	60	
QY	197	GGGACGTGGCGGCAGAGCTGGGGAATGGCGCGGAGCCAGCGCGGCTTCTCTCTCT	256	
Db	61	AGCACCGCGGTCGAGCGCGGGAATGGCGCGGAGCCAGCCCGCTCTTCTCTCT	120	
QY	257	GATGCTTTTCAGACCGCGGTCTCTGACGGGGCGCTTGTCTGCTTTCACACTCTCTGCACTA	316	
Db	121	GATCTCTCTGGGACCGTGGCGCGGAGCGGCGGACCCACCGGTCACTCTCTGCGCTA	180	
QY	317	CCTCTTCATGGGTGCTCAGAGCAGACCTTGGTCTTTCCTTGTGTTGAAGCTTTGGGCTA	376	
Db	181	CCTCTTCATGGGGCGCTCAGAGCGGGACACCGCCCTGCCCTGTGTTGAAGCCCTTGGGCTA	240	
QY	377	CGTGGATGACCACTGTTTCGTGTCTTATGATCATGAGATCGCGGTGTGAGCCCGCAAC	436	
Db	241	TGTGGATGACGAGCTTTTCGTGGCCCTACATCAGCAGAGTCGCGGTGCAGAGTCCCGTGC	300	
QY	437	TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGACGTGAGTCAGAGTCTGAA	496	
Db	301	CCAGTGGGTCCTGGGTGAGGCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCCTGAA	360	
QY	497	AGGGTGGGATCACAATGTTCACTGTTGACTTCTGGACTATTATGGAAATCACAACCAAC	556	

Db	361	AGGTGGGATCACATGTTTCATCGTCGACTCTTGACCATCATGTGACACCAACCAACACAG	420
QY	557	CAAGGAGTCCACACACCCCTGCAAGGTATCTCTGGGTGTGAAATGCAAGAAGACAACAGTAC	616
Db	421	CAAGGAGTCCACACACCCCTGCAAGGTATCTCTGGGTGTGAAATGCAAGAAGACAACAGTAC	480
QY	617	CGAGGGTACTGGAAGTACAGGGTATGATGGCAGGACCACTTGAATCTGCCCTGACAC	676
Db	481	CAGGGGTTCTGGAAGTACAGGGTACGATGGCAGGACCACTTGAATCTGCCCTGAGAC	540
QY	677	ACTGATTGGAGACGACAGAACACAGGGCTGGCCACCAAGCTGGAGTGGGAAGGCA	736
Db	541	ACTGATTGGAGACGACAGAGTCCAGGGCCCTGACCAACCAAGCTGGAGTGGGAAGTGA	600
QY	737	CAAGATTGGGCCAGGACAGAACAGGGCTACTCGAGAGGGACTGCCCTGCACAGCTGCA	796
Db	601	CAAGATTGGGCCAAGACAGACAGAGCTATCTCGAGAGGGACTGCCCTGAGCAGCTGCA	660
QY	797	GCAGTTGCTGAGCTGGGAGAGGTGTTTTGGACCAACAAGTGCCTCCTTTGGTGAAGGT	856
Db	661	GTGTTGCTGGAGCTGGGAGAGGGTCTTGACACAGCAAGTGCCTCTTTGGTGAAGGT	720
QY	857	GACACATCATGTGACCTTTCAGTGACCACTCTACGGTGTGGGCCCTTGAACCTACTACCC	916
Db	721	GACTCATCATGTGGCTCTGCAAGTACCACTCTACGGTGTGAGCTCTGAACCTTCTACCC	780
QY	917	CCAGAACATCACATGAAGTGGCTGAAGGATAAGCAGGCAATGGATGCAAGGAGTTCGA	976
Db	781	CCAGAACATCACATGAGTGGTGTGAAGGACAGAGGCAATGGATGCAAGGAGCGGA	840
QY	977	ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAGGGCTTGATTAACCTTTGGCTGT	1036
Db	841	GTCTAAGGATGTGCTGCCACGGGAGCGGACCTACCAGAGCTGGGAGGCTTTGGCCGT	900
QY	1037	ACCCCTGGGAAAGACAGACAGATATAGTGCAGGTGGAGCACCAGCCCTGGATCAGCC	1096
Db	901	GCCCCCTGGGAGAGCAGAGATATACCTGCAGGTGGAGCACCAGCCCTGGATCAGCC	960
QY	1097	CCTCATTTGTATCTGGAGCCCTCACCGCTGCGCACCTAGTCAATGGAGTCATCAGTGG	1156
Db	961	ACTCACTGCCACCTGGAGCCCTCAGTTCTTAACACCTCGTCACTGAGTCATCAGTGG	1020
QY	1157	AATTGCTGTTTTGTCTGTCATCTTTGTTCAATGGAAATTTGTTCAATAATTAAGGAAGAG	1216
Db	1021	GATTGCTGTTTGTCTATCATCTTCTTATTGGAAATTTGTTCAAGATCTTTAAGGAAAG	1080
QY	1217	GCAGGGTTCAGAGAGGCCATGGGCACTAGTCTTAGCTCAAGCTGAGTGACACGAGC	1276
Db	1081	ACAAGCTTCAAGAGAGCCATGGGGACTAGTCTTAGTGAATGTGAGTGACACGACAGC	1140
QY	1277	CTGCAGACTACTCTGGGAAGGACAAAACTACAGACTCAAGCAGGAGTGCATTATG	1336
Db	1141	CTGCAGACTCCTGTGGGAGGACACATATTTGGAGACTCAGAGGGGTGCAC--TTGGGG	1199
QY	1337	AGCTCTTCATGTTTTTCAGGAGAGTGTGAACCTAA--ACATAGAAATTCCTGACCACTCC	1395
Db	1200	CATTCTTTGTGTTTTCAGGACAGAGTTGAACCTACAATGGAACTGTCCAGAGGTCT	1259
QY	1396	TTGATTTTAGCCTTCTCTGTTTCATTTCTCTCAA	1428
Db	1260	CTGCTCTTAGCCTTCCCTGTTCACTCTCAAAA	1292

RESULT 14

AY007543

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AY007543

Dicerorhinus sumatrensis

AY007543

AY007543.1

GI:10945691

Sumatran rhinoceros.

Dicerorhinus sumatrensis

1320 bp

mRNA

linear

MAM 22-OCT-2000

HFE protein mRNA, complete cds.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus. 1 (bases 1 to 1320)
AUTHORS West,C.J., Worley,M. and Beutler,E.
TITLE Rhinoceros HFE Polymorphisms
JOURNAL Unpublished
2 (bases 1 to 1320)
AUTHORS West,C.J., Worley,M. and Beutler,E.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Molecular and Experimental Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
source Location/Qualifiers
1..1320
/organism="Dicerorhinus sumatrensis"
/db_xref="taxon:89632"
86..1132
/codon_start=1
/product="HFE protein"
/protein_id="AAG23703.1"
/db_xref="GI:10945692"
/translation="MGPRAPALFELILRTVAAGRPPRSHSLRYLFMGASRDRGLPLFEALGYVDDELFAVYNHESRAESPAQVNLGAEHSQWLQLSQSLKGNDHMFIVDFRLTMDNNHNSKHTQVLIGCEVQEDRSCPEQWLMLGELRGYDQVPPVLKVVTHHVASRAITLKEVNRKIRAKQNRLEYRDCPEQLQWLELGRDLPSPGDTQSWVALAVPPEAVYTLRCQALNFYQNTIRMLKDRKPDVKDAESKDLPSGDGTQSWVALAVPPEEORTYCGVHGPHLDQPLTATWEPSSLNTLVGTVISGIAVCVIFILIGILFRILKRQQA SRGAGDXYLAEEC"

BASE COUNT	281 a	358 c	398 g	283 t
ORIGIN				
Query Match	62.3%;	Score 897;	DB 4;	Length 1320;
Best Local Similarity	82.4%;	Pred. No. 4.5e-252;		
Matches 1065; Conservative	0;	Mismatches 225;	Indels 3;	Gaps 3;

QY 137 AGGGTGACTTCTTGAGGCATCCCGGTTTCCCGCGCCCGCCAAAGAGCGGAGTTTAAACG 196
Db 2 AGGGTGACTTCTTGGGTCACTCCTTATTTCCCGCCCGCCAGCAGCAGTGCACA -CGGSCC 60
QY 197 GGACGTGGGCCAGAGCTGGGGAATATGGCCCGCGAGCCAGCGCGGCTTCTCTCCT 256
Db 61 AGCACCCCGCGTCCGAGCCGGGGAATATGGGCGCGAGCCGCGCTGTCTCTCCT 120
QY 257 GATCGTTTGCAGACCGCGTCTCGTAGGGCGCTTGTGCGTTTACACTCTCTGCCTA 316
Db 121 GATCCTCTCGGAACCTGGCCGCGCAGGGCGAGCGCGCGTGCACACTCTCGCGCTA 180
QY 317 CCTTTCATGGTGCCTCAGACGAGCACTTGGTCTTTTCTTGTGTAAGCTTTGGGCTA 376
Db 181 CCTTTCATGGCGCCTCAGACGGGACCAGCGCTGCCCCCTGTTTGAAGCTTTGGGCTA 240
QY 377 CGTGGATCACCAGTGTTCGTTCATGATCATGAGTCCCGCTGGAGCCCCGAC 436
Db 241 TGTGGATGACGAGCTTTTCGTGGCGTACAATCATGAGATGCGCGTCGAGAGTCCCCTGC 300
QY 437 TCCATGGTTTCCAGTAGAATTTCAAGCCAGATGTGCGCTGCAGCTGAGTCAGACTCAA 496
Db 301 CCAGTGGGTCTTGGGTGAGGCCCAACAGCCAGCTGTGGCTGCAGCTGAGTCAGACCTGAA 360
QY 497 AGGTGGGATCACATGTTCACTGTTCGACTTCTGGCACTATTATGGAATACACACACACAG 556
Db 361 AGGTGGGATCACATGTTTCATCTCGACTTCTGGACCATATGACACACACACACACAG 420
QY 557 CAAGAGTCCACACCCCTGCAGGTATCCTGGGCTGTGAATGCAAGAAGACACAGTAC 616
Db 421 CAAGAGTCCACACCCCTGCAGGTATCCTGGGCTGTGAGTGTGAGGAGGACACACAGC 480
QY 617 CGAGGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTGAATCTGCCCTGAC 676
Db 481 CAGGGGTCTTGGAAGTACGGGTACGATGGCAGGACCATCTTGAATCTGCCCTGAGC 540
QY 677 ACTTGATTTGAGAGCAGCAAGAACCCAGGGCTTGCCCAACCAAGCTGGAGTGGGAAGCA 736


```

Db 721 GACTCATCATGTGGCCCTCTGCAGTGACCACTCTACGGTCTCAGGCTCTGAACCTTCTACCC 780
QY 917 CCAGAAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAAATGGATGCCAAGGAGTTGGA 976
Db 781 CCAGAACATCACCATGAGTGGTGAAGGACAGGAAGCCAGTGGATGTCAAGGACGCCGA 840
QY 977 ACCTAAGACGTATTCGCCAATGGGATGGGACCTACCAAGGCTGGATAACCTTGGCTGT 1036
Db 841 GTCTAAGGATGTCTCCCGCAGGGGATGGGACCTACCAAGAGCTGGGTGGCTTTGGCTGT 900
QY 1037 ACCCCCTGGGGAGACAGAGATATACGTGCCAGTGGAGCACCCAGGCCCTGGATCAGCC 1096
Db 901 GCCCCCGGGGAGAGACAGATATACCTGCCAGGTGGAGCACCCAGGCCCTGGATCAGCC 960
QY 1097 CCTCATTTGTATCTGGAGCCCTCACCGTCTGGCACCCCTAGTCAATGGAGTCATCAGTGG 1156
Db 961 CCTCACTGCCACCTGGGAGCCCTCACTTCTAACACCCCTCGTCACTGGAGTCATCAGTGG 1020
QY 1157 AATTGCTGTTTTGTCGTCACTCTGTTCATTTGGAATTTTGTTCATAATAATTAAGGAAGAG 1216
Db 1021 GATTGCTGTTGTGTCATCATCTTCTCATTTGGAATTTTGTTCAGAACTTTTAAGGAAAAG 1080
QY 1217 GCAGGTTCAAGAGAGCCATGGGCACTACGTCTTAGCTGACCTGAGTGAACGACGACG 1276
Db 1081 ACAAGCTTCAAGAGAGCCATGGGGACCTACGCTTTAGCTGAATGTAGTGAAGCACAGC 1140
QY 1277 CTGCAGACTCACGTGTGGGAGGAGACAAAACCTAGAGACTCAAAGAGGGAGTGCATTATG 1336
Db 1141 CTGCAGACTCCTTGTGGGAGGACACATAATTGGAGACTCAGAGAGGTGCAC-TGGTGG 1199
QY 1337 AGCTCTTCATGTTTCAGGAGAGAGTTGAACCTTAACATAGAAATGCGCTGACGAACTCCT 1396
Db 1200 CGTTCTTCGTGTTTCAGGACAGAGTTGAACCTTAACAAATGGAACTGTCCCAGAGGTCT 1259
QY 1397 TGATTTTAGCCTTCTCTGTTCATTTCTCTCAA 1428
Db 1260 CTCTCTTAGCCTTCCCTGTTCACCTCTCAAAA 1291
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:34 ; Search time 2546.54 Seconds
(without alignments)
7632.166 Million cell updates/sec

Title: US-09-497-957-9

Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAGATTCCCCA 1440

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	738.6	51.3	819	10	BG747345
2	633.8	44.0	1723	11	AK009581 Mus muscu
3	506.8	35.2	570	10	BE272926
4	358	24.9	384	10	BF863952
5	284.2	19.7	550	10	BI339179 PM4-ET020
6	276.6	19.2	523	10	BF080089
7	259.4	18.0	489	10	BE994943
8	251.4	17.5	457	9	AI850020
9	248.4	17.2	455	10	BE995172
10	231.6	16.1	464	9	AA217236
11	226.4	15.7	268	10	W21141
12	214.6	14.9	546	10	BE809138
13	211.2	14.7	481	9	BB851691
14	203.8	14.2	831	10	BI452668
15	198.8	13.8	502	9	BB851691
16	198.2	13.8	392	10	BF465475
17	187.4	13.0	1208	11	BC005306 Homo sapi

18	174	12.1	394	10	BF464345
19	170.4	11.8	444	12	AZ025590
20	160	11.1	334	9	AW902003
21	156.8	10.9	536	12	AZ074871
22	154	10.7	805	10	BG402460
23	153.6	10.7	1016	9	AL532717
24	151.8	10.5	289	10	H33644
25	151.6	10.5	481	12	AZ025784
26	146	10.1	531	12	AQ703007
27	145.6	10.1	1031	9	AL513728
28	143.6	10.0	1298	11	AK005051
29	143.2	9.9	1121	9	AL545738
30	141.6	9.8	972	9	AL541209
31	141.4	9.8	959	9	AL560013
32	140.8	9.8	942	9	AL539924
33	139.8	9.7	910	9	AL558822
34	139.4	9.7	1020	9	AL535151
35	138	9.6	1094	10	BM465674
36	134.2	9.3	932	9	AL547313
37	133.8	9.3	936	10	BI832514
38	133	9.2	679	10	BG680802
39	132.6	9.2	937	9	AL540488
40	132.6	9.2	1090	10	BM477276
41	131.6	9.1	930	9	AL541347
42	131.6	9.1	979	9	AL542403
43	131.6	9.1	979	9	AL557373
44	131.4	9.1	538	9	BE232472
45	131.2	9.1	916	10	BG744157

ALIGNMENTS

RESULT 1

BG747345

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG747345 819 bp mRNA linear EST 15-MAY-2001
602704818F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4857941 5',
mRNA sequence.
BG747345 GI:14057998
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 819)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: AFCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CMI711 row: d column: 06

High quality sequence stop: 792.

Location/Qualifiers

1..819

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4857941"

/clone_lib="NIH_MGC_15"

/tissue_type="adenocarcinoma cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pOTB7; Site:1: XhoI; Site_2:

ECORI; cDNA made by oligo-dt priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)"									
BASE COUNT	202 a	201 c	235 g	181 t					
ORIGIN									
Query Match	51.3%	Score	738.6:	DB 10:	Length	819;			
Best Local Similarity	96.1%	Pred	No. 2.2e-194;						
Matches	789;	Conservative	0;	Mismatches	29;	Indels	3;	Gaps	3;
QY	384	GACCAGCTGTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCGAATCCATGG	443						
Db	1	GACCAGCTGTCGTGTTCTATGATCATGAGAGTCGCCGTGTGGAGCCCGAATCCATGG	60						
QY	444	GTTCACAGTAGAATTCACAGCCAGATGTGGCTGCAGCTCAGTGCAGAGTCTGAAAGGGTGG	503						
Db	61	GTTCACAGTAGAATTCACAGCCAGATGTGGCTGCAGCTCAGTGCAGAGTCTGAAAGGGTGG	120						
QY	504	GATCACATGTTCACTGTTGACTTCTGCACTATTATGGAATATCACACACACGACGAGGAG	563						
Db	121	GATCACATGTTCACTGTTGACTTCTGCACTATTATGGAATATCACACACACGACGAGGAG	180						
QY	564	TCCACACCCCTGCAGTCACTCTGGGCTGTGAATGCCAAGACACAACTACCGAGGCG	623						
Db	181	TCCACACCCCTGCAGTCACTCTGGGCTGTGAATGCCAAGACACAACTACCGAGGCG	240						
QY	624	TACTGGAAGTACCGGTATGATGGCAGGACCACCTTGAATCTGCCCTGCACACTGGAT	683						
Db	241	TACTGGAAGTACCGGTATGATGGCAGGACCACCTTGAATCTGCCCTGCACACTGGAT	300						
QY	684	TGGAGAGCAGCAGAACCCAGGGCCTGCCCCACCAAGCTGGAGTGGGAAGGCACAGATT	743						
Db	301	TGGAGAGCAGCAGAACCCCA-GGCCTGGCCCAACCAAGCTGGAGTGGGAAGGCACAGATT	359						
QY	744	CGGGCCAGGCAGACAGGGGCTACCTGGAGAGGGAGTCCCTGCACAGCTGCAGCAGTTG	803						
Db	360	CGGGCCAGGCAGACAGGGGCTACCTGGAGAGGGAGTCCCTGCACAGCTGCAGCAGTTG	419						
QY	804	CTGGAGCTGGGGAGAGTGTTTGGACCACCAAGTGCCTCTTTGGTGAAGTGCACACAT	863						
Db	420	CTGGAGCTGGGGAGAGTGTTTGGACCACCAAGTGCCTCTTTGGTGAAGTGCACACAT	479						
QY	864	CATGTGACCTCTTCAGTGACCACTCTACGGTGTGGGGCCTTGAACCTACTACCCCGAGAC	923						
Db	480	CATGTGACCTCTTCAGTGACCACTCTACGGTGTGGGGCCTTGAACCTACTACCCCGAGAC	539						
QY	924	ATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGGATGCCAAGGAGTTCGAACCTAAA	983						
Db	540	ATCACCATGAAGTGGCTGAAGGATAGAGCAG-CAATGGATGCCAAGGAGTTCGAACCTAAA	598						
QY	984	GACGTATTGCCCAATGGGGATGGGACCTACCAAGGGCTGGATAACCTTGGCTGTACCCCTT	1043						
Db	599	GACGTATTGCCCAATGGGGATGGGACCTACCA-GGCTGGGATAACCTTGGCTGTACCCCTT	657						
QY	1044	GGGGAAGCAGAGATATACCTGCCAGGTGGAGCAGCCAGGCTGGATCAGCCCTCAT	1103						
Db	658	GGGGAAGCAGAGATATACCTGCCAGGTGGAGCAGCCAGGCTGGATCAGCCCTCAT	717						
QY	1104	GTGATCTGGGAGCCCTCACCCTGTGGCACCCTAGTCATTTGGAGTCACTAGTGAATTGCT	1163						
Db	718	GTGATCTGGGAGCCCTCAGCTCTGGCACCTAGTCATTTGGAGTCACTCCAGTGGAAATTCG	777						
QY	1164	GTTTTTCGTGTCATCTTGTTCATTTGGAAATTTGTTGCATAAT	1204						
Db	778	TGTTTTTCGTGTCATCTTGTTCATTTGGAAATTTGTTGCATAAT	818						
RESULT	2								
AK009581									
LOCUS									
DEFINITION									

ACCESSION	AK009581.1	GI:12844462
VERSION	HTC: CAP trapper.	
KEYWORDS	Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,	
SOURCE	clone_lib:RIKEN full-length enriched mouse cDNA library	
	clone:2310032M04.	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (sites)	
AUTHORS	Carninci,P. and Hayashizaki,Y.	
TITLE	High-efficiency full-length cDNA cloning	
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE	99279253	
PUBMED	10349636	
REFERENCE	2 (sites)	
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE	20499374	
PUBMED	11042159	
REFERENCE	3 (sites)	
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Fujimoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Harashi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.	
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE	20530913	
PUBMED	11076861	
REFERENCE	4 (sites)	
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
TITLE	Functional annotation of a full-length mouse cDNA collection	
JOURNAL	Nature 409, 685-690 (2001)	
REFERENCE	5 (bases 1 to 1723)	
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yanaiishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome-gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.	
	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAAGGACCAAGAGCTCTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase	


```

Db 324 ACCAGCTG-TCTGTCTATGATCATGAGTGCCTGTGGAGCCCGAATCCATGGG 266
QY 445 TTTCAGTAGAATTCAAGCCAGATGTGCTCAGCTCAGTCTGAAAGGTTGG 504
Db 265 TTTCAGTAGAATTCAAGCCAGATGTGCTCAGCTCAGTCTGAAAGGTTGG 206
QY 505 ATCAGATGTTCACTGTGACTTCTGGACTTATTTATGAAATATCACAACACAGCAGGAGT 564
Db 205 ATCAGATGTTCACTGTGACTTCTGGACTTATTTATGAAATATCACAACACAGCAGGAGT 146
QY 565 CCCACACCTGCAGGTCATCTCTGGCTGTGAAATCAAGAACACACAGTACCGAGGCT 624
Db 145 CCCACACCTGCAGGTCATCTCTGGCTGTGAAATCAAGAACACACAGTACCGAGGCT 86
QY 625 ACTGGAAGTACGGGTATGATGGGAGGACCACTTGAATTCGCCCTGACACACTGGATT 684
Db 85 ACTGGAAGTACGGGTATGATGGGAGGACCACTTGAATTCGCCCTGACACACTGGATT 26
QY 685 GGAGAGCAGC 694
Db 25 GGAGAGCAGC 16

RESULT 5
BI339179
LOCUS 364041 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 30-JUL-2001
DEFINITION BI339179
ACCESSION BI339179
VERSION BI339179.1 GI:15032462
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 550)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCAGCAGC
Plate: 100 row, C column: 24
Seq primer: ATTTAGTGACACTATAG.
Location/Qualifiers
1. .550
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 108 a 180 c 164 g 98 t
ORIGIN

Query Match 19.7%; Score 284.2; DB 10; Length 550;
Best Local Similarity 74.1%; Pred. No. 5.8e-68;
Matches 413; Conservative 0; Mismatches 113; Indels 31; Gaps 3;

```

```

QY 140 GTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAGACGCGAGATTTTAAACGGG 199
Db 1 GTGACTTCTGGAGCCCTCGGTTTCCCGCCCCCAACACGCGCCGAAAAG-----CCCTGG 55
QY 200 ACCTGGCCAGAGCTGGGAAATGGGCCCGCCGAGCCAGCCCGCGCTTCTCCTCCTGAT 259
Db 56 AAGCGCGCTCGAGCCGG--AAAGGGCCCAACAAGCCCGCGCGCTTCTCCTCCTGAT 113
QY 260 GCTTTTTCAGACCCCGCTCTCGAAGGGCGCTTGCCTGGTTTCACACTCTCGCACACT 319
Db 114 CCTCCTCGGACCTGGCCAGCAGAGCGAGCCCGCCGACACTCCCTGCTCTTCT 173
QY 320 CTTCATGGGTGCTTCAGAGCAGGACCTTGGTCTTCTTCTTCTTGAAGCTTTGGGCTACGT 379
Db 174 CTTCATGGGCTCGGAGCCAGATCTCGGCTGCCCTCTTTTCAGGCTTTGGGCTACGT 233
QY 380 GGATGACCACTGTTCTGTTCTATGATGAGAGTCGCGCTGGAGCCCGCAACTCC 439
Db 234 GGAGCACCAGCTGTTCTGTCTACAATCACGAGAGTCGCGCTCGAGAGCTCGCGCCC 293
QY 440 ATGGTTTCCAGTAGAATTTCAAGCCAGATGTGCTGCAGCTGAGTCAGAGTCTGAAAG 499
Db 294 CTGGCTCTCGGTAAGGCTTCCACACCACTGGCTGCAGCTAAGCCAGAGCTGAAAG 353
QY 500 GTGGATCACATGTTCACTTCTGACTTCTGACTTATTTATGAAAANTCACAACACAGCAA 559
Db 354 GTGGATCACATGTTCACTTCTGACTTCTGACTTCTGACTTCTGAGATGAGTGGGAGGCGCAGACCA 413
QY 560 G-----GAGTCCCACACCCCTGCAGGTCATCTCGGCTGTGA 595
Db 414 GATTAACCAACTGGGAGTGTGCCAGAGTCCCACACCCCTGCAGGTCATCTCGGCTGTGA 473
QY 596 AATCAAGAAAGACAACAGTACCAGGCTACTGGAAGTACGGGTATGATGGGCGAGCACA 655
Db 474 AGTCAAGCGGACAACAGCAGCAGAGGTTCTGGAATGATGAGTGGGAGGCGCAGACCA 533
QY 656 CTTGTAATTCGCCCTG 672
Db 534 CTTGGAGTTCCACCCCTG 550

RESULT 6
BF080089
LOCUS 230846 MARC 2P1G Sus scrofa cDNA 5', mRNA linear EST 18-OCT-2000
DEFINITION BF080089
ACCESSION BF080089
VERSION BF080089.1 GI:10873919
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 523)
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCAGCAGC
Plate: 48 row: E column: 9
Seq primer: ATTTAGTGACACTATAG.

```

COMMENT

Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643 USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
The following repetitive elements were found in this cDNA sequence:
1-31. >(CAG)n#SimpleRepeat
Seq primer: M13 Forward
POLYA-No.

FEATURES

Location/Qualifiers
source
1..489
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-CG0p-bik-d-03-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73B-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP_Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu.
TAG_SEQ=None found"

BASE COUNT

103 a 133 c 144 g 109 t

Query Match

Best Local Similarity 18.0%; Score 259.4; DB 10; Length 489;
Matches 357; Conservative 0; Mismatches 106; Indels 24; Gaps 1;

Qy 238 GCGCGCGCTTCTCCTCGTATGCTTTTCAGACGCGGGTCTTCAGGGGGCTTGTGTC 297
Db 3 GCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 62
Qy 298 GTTCACACTCTGACACTACCTCTTCATGGGTGCTTCAGACGAGACCTTGGTCTTTCCT 357
Db 63 GTTCACATCTCTAGATACCTCTTCATGGGTGCTTCAGACGAGACCTTCGGCTGCTT 122
Qy 358 TGTGTAAGCTTTGGGCTACGTGGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
Db 123 TGTGTAAGCTTTGGGCTACGTGGATGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
Qy 418 GCGGTGTGAGCGCCGAACTCCATGGGTTCACAGTAGAATTTCAAGCCAGATGTGGTGTG 477
Db 183 GCGGTGTGAGCGCCGAACTCCATGGGTTCACAGTAGAATTTCAAGCCAGATGTGGTGTG 242
Qy 478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATACATGTTTCACTGTTGACTTCTGGACTATTA 537
Db 243 ATCTGAGTCAGAGCTGAAAGGGTGGGATACATGTTTCACTGTTGACTTCTGGACTATTA 302
Qy 538 TGGAAATACAAACACACACAG-----CAGTCCCAACACCC 573
Db 303 TGGCAACTATAACACACAGTAAAGTTCAGAAAGTGGGAGTGGTTCAGAGTCCCAACACCC 362
Qy 574 TCGAGTCTATCTGCGCTGTGAAATGCAAGACACAGTACCGAGGCTACTGGAAGT 633
Db 363 TCGAGTCTATCTGCGCTGTGAAATGCAAGACACAGTACCGAGGCTACTGGAAGT 422
Qy 634 ACGGTTATGATGGCAGGACACCTTGAATTTGCCCTGCACACACTGGATTTGGAGAGCAG 693
Db 423 ATGGTTATGATGGCAGGACACCTTGAATTTGCCCTGCACACACTGGATTTGGAGAGCAG 482

FEATURES

Location/Qualifiers
source
1..523
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
103 a 175 c 152 g 93 t

Query Match 19.2%; Score 276.6; DB 10; Length 523;
Best Local Similarity 74.2%; Pred. No. 7.4e-66;
Matches 388; Conservative 0; Mismatches 109; Indels 26; Gaps 2;

Qy 119 AGCTAGCAATAGCTGTAGGTGACTTCTGGAGCATCCCGTTTCCCGCCGCCCAAA 178
Db 2 AGCTTGGCAATGGCTCCAGGGTGACTTCTTGAGCCACCTCGGTTTCCCGCCGCCCAAA 61
Qy 179 AGAAGGGGAGATTACGGGGACCTCGGCCAGAGCTGGGAAATGGGCGCGGAGCCAG 238
Db 62 AGCGCGCGAAAGACACCTCGAAGCGCGGTCCGAGCC--GGGAAATGGGCGCGGAGCCG 119
Qy 239 GCGCGGCTTCTCCTCTGATGCTTTTCAGACGCGGGTCTTCAGAGGGCGCTTGTGCG 298
Db 120 GCGGGGCTTCTCCTCTGATGCTTTTCAGACGCGGGTCTTCAGAGGGCGCTTGTGCG 179
Qy 299 TTACACTCTCTGACTACCTCTTCATGAGGTGCTTCAGAGGAGCTTGTGCTTTCCTT 358
Db 180 GCCACACTCCTCTCTCTCTTCATGAGGCGCTCGGAGCGAGATCTCGGCTGCGCCCT 239
Qy 359 GTTTGAGCTTTGGGCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 418
Db 240 GTTTGAGGCTTTGGGCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 299
Qy 419 CCGTGTGGAGCCCGCAACTCCATGGGTTCAGAGTAAATTTCAAGCCAGATGTGGGTGCA 478
Db 300 CCGTGTGGAGCCCGCAACTCCATGGGTTCAGAGTAAATTTCAAGCCAGATGTGGGTGCA 359
Qy 479 GCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTTCACTGTTGACTTCTGGACTATTA 538
Db 360 GCTAAGCCAGAGCTGAAAGGGTGGGATCACATGTTTCACTGTTGACTTCTGGACTATTA 419
Qy 539 GGAATATCAACACACAGCAAG-----GAGTCCCAACACCT 574
Db 420 GGAATATCAACACACAGCAAG-----GAGTCCCAACACCT 479
Qy 575 GCAGGTATCTCTGGGCTGTGAAATGCAAGAGACACAGTACC 617
Db 480 GCAGGTATCTCTGGGCTGTGAAATGCAAGAGAGAGTGTGGCCAGAGTCCCAACACCT 522

RESULT

BE994943 489 bp mRNA linear EST 05-OCT-2000
LOCUS
DEFINITION
UT-M-CG0p-bik-d-03-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
UT-M-CG0p-bik-d-03-0-UI 3', mRNA sequence.
ACCESSION
BE994943
VERSION
BE994943.1 GI:10678689
KEYWORDS
EST.
SOURCE
house mouse.
MUS MUSCULUS
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 489)
REFERENCE
AUTHORS
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477

QY 694 CAGAACC 700
 Db 483 CCGAGCC 489

RESULT 8

AI850020

LOCUS AI850020 457 bp mRNA linear EST 15-JUL-1999
 DEFINITION UI-M-BGO-aib-g-10-0-UI.s1 NIH_BMAP_MSC Mus musculus cDNA clone

UI-M-BGO-aib-g-10-0-UI 3', mRNA sequence.

AI850020

AI850020.1 GI:5493926

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 457)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

Oligo-dt track not found, Not I site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.

Soares Lab Clone distribution: NIH BMAP cDNA clones will be made

available by the means that is soon to be determined. When NIH

determines the means for distribution of the BMAP cDNA clones, this

record will be updated accordingly when that means is determined.

The following repetitive elements were found in this cDNA sequence:

3-30, >(CAG)n\$Simple.repeat

Seq primer: M13 Forward

POLYA-No.

Location/Qualifiers

1..457

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-BGO-aib-g-10-0-UI"

/dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site.1: Not I; Site.2: Eco RI; The

NIH_BMAP_MSC library is a non-normalized library

constructed from mouse spinal cord. The tag is a string

of 5 nucleotides present between the Not I site and the

oligo-dt track. The library was constructed as described

by Bonaldo, Lennon and Soares, Genome Research 6: 791-806

, 1996. Tissue provided by Ms. Annie Novakovich,

Zivic-Miller Laboratories.

TAG_LIB=NIH_BMAP_MSC

TAG_TISSUE=spinal-cord

TAG_SEQ=TCGAA"

BASE COUNT 100 a 129 c 144 g 84 t

ORIGIN

Query Match 17.5%; Score 251.4; DB 9; Length 457;

Best Local Similarity 79.6%; Pred. No. 7.2e-59;

Matches 297; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 561 GAGTCCACACCTCGAGGTATCTGGGCTGTGAATGCAAGAGACAACAGTACCAG 620

Db 85 GAGTCCACATCTCGAGGTGTCTAGGCTGTGAGGTGATGAAGACAACAGTACCAGC 144

QY 621 GGCTACTGGAAGTACGGGTATGATGGGCAGCAGCACCTTGAATTCGGCTGCACACTG 680
 Db 145 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTGGAAATTCGCCCAACACTA 204
 QY 681 GATTGGAGAGCAGCAGAACCCAGGCGCTGGCCCAACAGCTGGAGTGGGAAAGGCACAAG 740
 Db 205 AACTGGAGCGCAGCGAGCGGCTGGGCGCACCAGGTGAATGGAGCGAGCACAAG 264
 QY 741 ATTGGGCGCAGCAGAACAGGCGCTACCTGAGAGGAGTCCCTGCACAGCTGCAGCAG 800
 Db 265 ATCCGTGCCAACAGACAGGAGTACCTGGAGAGGAGTCCCGCCGAGCAGCTGAAACGG 324
 QY 801 TTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGCCTCTTGGTGAAGGTGACA 860
 Db 325 TCCTTGGAGCTGGGAGAGGCGTTCTGGCAGCAGCAAGTGCCTACTTTGGTGAAGTGNCT 384
 QY 861 CATCATGTGACCTCTTCACTAGTACCACACTACAGGTGTGGGGCTTGAACACTACTACCCAG 920
 Db 385 CGCCACTGGGCTCTACGGGAGCCTCTCTAAGGTGTGAGGCTCTGGACTTCTTCCCGCAG 444
 QY 921 AACATCACCATGA 933
 Db 445 AACATCACTATGA 457

RESULT 9

BE995172

LOCUS

DEFINITION

UI-M-CG0p-bil-h-10-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone

UI-M-CG0p-bil-h-10-0-UI 3', mRNA sequence.

BE995172

BE995172.1 GI:10679153

EST.

KEYWORDS

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 455)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

Oligo-dt track not found, Not I site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.

Soares Lab Clone distribution: Researchers may obtain BMAP cDNA

clones from RESEARCH GENETICS. It should be noted that Bento Soares

is generating a small number of additional specialized

non-redundant arrays of BMAP cDNAs whose availability will be

considered under appropriate and limited collaborative arrangements

The following repetitive elements were found in this cDNA sequence:

1-31, >(CAG)n\$Simple.repeat

Seq primer: M13 Forward

POLYA-No.

Location/Qualifiers

1..455

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UI-M-CG0p-bil-h-10-0-UI"

/clone_lib="NIH_BMAP_Ret4_S2"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site.1: Not I; Site.2: Eco RI; The

NIH_BMAP_Ret4_S2 library is a subtracted library.

FEATURES

source

ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu.

BASE COUNT	99 a	129 c	144 g	83 t
ORIGIN				

Query Match	17.28;	Score	248.4;	DB	10;	Length	455;
Best Local Similarity	79.5%;	Pred. No.	4.9e-58;				
Matches	294;	Conservative	0;	Mismatches	76;	Indels	0;
Gaps	0;						
QY	561	GAGTCCACACCCCTGCAGTCTATCTGGCGTGCAAATGCAAGAAGACAACTGATCCGAG	620				
DB							
QY	86	GATCCCAACATCTCGAGTGGTCTTAGGCTGTGAGTGGCATGAGACACACATGATCCAGC	145				
DB							
QY	621	GGCTACTGGAATACGGGTATGATGGCGAGACACACCTTGAATCTGCGCTGCACACTG	680				
DB							
QY	146	GGCTCTTGAGATATGGTTATGACGGCGAAGATCACCTGGAAATCTGCCCAAGACACTA	205				
DB							
QY	581	GATTGAGAGCAGCAGACACCCAGCGCCTGGCCCAACCAAGCTGGAGTGGGAAGGCAACAG	740				
DB							
QY	206	AAGTGGAGCGCCACCGCGCCAGGGCCCTGGGCCCAACCAAGGTGGAATGGGACGAGCAAG	265				
DB							
QY	741	ATTTCGGGCCAGGCGAACAACAGGCGCTTACCTGGAGAGGAGCTGCCCTGCACAGCTGCAGCAG	800				
DB							
QY	266	ATCCGTGCCAACAGACAGAGGAGCTTACCTGGAGAAGGACTGCCCGAGCAGCTGAACGG	325				
DB							
QY	801	TTGCTCGAGCTGGGAGAGGTGTTTGGACCAACAAGTGCCTCCTTTGTTGAAGGTGACA	860				
DB							
QY	326	CTCCTCGAGCTGGGAGAGGCCCTTCTGGGCAGCAAGTGCCTACTTTTGGTGAAGTGACT	385				
DB							
QY	861	CATCATGTGACCTTTCAGTGACCACCTACGCGTGTGGGCTTTGAAGTACTACTACCCCCAG	920				
DB							
QY	386	CGCCTCTGGCCCTTACGGGGACCTCTCTTAGGTGTGAGGCTCTGAGCTTCTTCCCCCAG	445				
DB							
QY	921	AACATCACCA	930				
DB							
QY	446	AACATCACTA	455				
DB							

RESULT 10	AA217236/c	LOCUS	AA217236	464 bp	mRNA	linear	EST 06-FEB-1997
DEFINITION			mu89005.r1	Soares mouse lymph node	NBMLN	Mus musculus	cDNA clone
			IMAGE:652689	5', similar to TR:G940354	G940354	CLASS I	
				HISTOCOMPATIBILITY ANTIGEN-LIKE PROTEIN. .: mRNA sequence.			

ACCESSION	AA217236	
VERSION	AA217236.1	GI:1826237
KEYWORDS	EST.	
SOURCE	mus mouse.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (bases 1 to 454)	
	Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kubacka, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellensberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, P.	

TITLE
 The WashU-HHMI Mouse EST Project
JOURNAL
 Unpublished (1996)
COMMENT
 Contact: Marra M/Mouse EST Project
 WashU-HHMI Mouse EST Project
 Washington University School of MedicineP
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:398537

Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham.

FEATURES
source

```

source
i. .404
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_image="IMAGE:652689"
/clone_lib="Soares mouse lymph node NDBMLN"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: lymph node; Vector: pT7T3D-P
with a modified polylinker; Site_1: Not I;
1st strand cDNA was primed with a Not I -
[5'
TCGTACCATTCGAAGTCGGAGCGCGCATCTTTT
3']; double-stranded cDNA was ligated to E
(Pharmacia), digested with Not I and clone
and Eco RI sites of the modified pT7T3 vec
provided by Dr. Bertrand Jordan. Library c
normalized by Bento Soares and M.Fatima Bo
101 a 136 c 119 g 108 t
BASE COUNT
ORIGIN

```

BASE COUNT	101 a	136 c	119 g	108 t
ORIGIN				

Query Match	16.1%	Score 231.6	DB 9	Length 464	
Best Local Similarity	74.3%	Pred. No. 2.4e-53			
Matches 330	Conservative	0	Mismatches 89	Indels 25	Gaps 2
Qy	294	CTGGGTTACAACTCTCTGCACTACTCTCTTCATGGTGCCTCAGAGCAGGACCTTGGTCTT	353		
Db	453	CCGGTTTCACATCTCTAAGATACCTCTTCATGGTGCCTCAGAGCAGCCTCGGGCTG	394		
Qy	354	TCCTTGTGTTGAAGCTTTGGGCTACGTGGATGACAGCTGTCGTGTCATGATCATGAG	413		
Db	393	CCTTTGTGTTGAGGCTAGGGGCTATGTGGATGACAGCTCTTGTGTCTCAACATCATGAG	334		
Qy	414	AGTCCCGTGTGGAGCCCGCAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGCG	473		
Db	333	AGTCCCGTGTCTGAGCCCGGGCCCGTGGATCTTGGAGCAAACTCCAAGCAGCTGTGG	274		
Qy	474	CTGCAGCTGAGTCAGAGTCTGAAGGGTGGATCACATGTCATGTTGACTTCTGGACT	533		
Db	273	CTGCATCTGATCAGAGCCTGAAGGGTGGAGTACATGTTTCATAGTAGACTTCTGGACC	214		
Qy	534	ATTATGGAATACAAACCACAGCAAG-----GAGTCCCCAC	569		
Db	213	ATCATGGCAACTATACCCAGTAAAGTCACGAAGTTGGAGTGGTCTCCGAGTCCCAC	154		
Qy	570	ACCTCGAGGTTCATCTCTGGGCTGTGAATGCAAGAACACACATACCGAGGGCTACTGG	629		
Db	153	ATCCTCGAGGTGTCTCTAGGCTGTGAGTGATGAAGACACACATACCGGGTCTTGG	94		
Qy	630	AAGTACGGGTATGATGGCAGGACCACTTGAATTTCTGCCCTGCACACTCGATTGGAGA	689		
Db	93	AGATATGGTTATACGGGCAAGATCACCTGGAATTTCTCCCAAGACACTAAACTGGAG-	33		
Qy	690	GCAGAGAACCCAGGGCTTGGCCC	713		
Db	34	GCAGCCGAGCCAGGGCTTGGCAC	11		

RESULT	11	
LOCUS	W21141	
DEFINITION	W21141 z851906.r1 Soares_fetal_lung_NbHL19 Homo sapiens cDNA clone IMAGE:307162 5', mRNA sequence.	linear EST 20-AUG-1996
ACCESSION	W21141	
VERSION	W21141.1	
KEYWORDS	EST.	
SOURCE	human.	

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 268)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
,M., Hultman,M., Kucabs,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478
Contact: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1642 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 173.
Location/Qualifiers
1..268
/organism="Homo sapiens"
/db_xref="GDB:1250574"
/db_xref="taxon:9606"
/clone_lib="IMAGE:307162"
/clone_lib="Soares_fetal_lung_NbHL19w"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: pVT3D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGCGCGCAATTTTTTTTTTTT-3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pVT3 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Patima Bonaldo. This library was constructed
from the same fetus as the fetal heart library. Soares
fetal heart NbHL19w."

BASE COUNT 79 a 55 c 64 g 69 t .1 others
ORIGIN

Query Match 15.7%; Score 226.4; DB 10; Length 268;
Best Local Similarity 99.2%; Pred. No. 5.1e-52;
Matches 238; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1201 TAATATTAGGAAGGAGGAGGTTCAAGAGGAGCCATGGGCACTTACGCTTAGCTGAAC 1260
|||||
Db 1 TAATATTAGGAAGGAGGAGGTTCAAGAGGAGCCATGGG-ACTACGCTTAGCTGAAC 59
|||||

QY 1261 GTGAGTGACACCGCCTGCAGACTCAGTGTGGNAGAGACAAACTAGAGACTCAAAG 1320
|||||

Db 60 GTGAGTGACACCGCCTGCAGACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAAG 119
|||||

QY 1321 AGGAGTGATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAT 1380
|||||

Db 120 AGGAGTGATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAT 179
|||||

QY 1381 TGCCTGACGAACCTCTTGATTTAGCCCTTCCTGTTTCATTTTCCTCAAAAGATTTCGCCA 1440
|||||

Db 180 TGCCTGACGAACCTCTTGATTTAGCCCTTCCTGTTTCATTTCTCTCAAAATAATTCGCCA 239
|||||

RESULT 12
BE809138
LOCUS BE809138 546 bp mRNA linear EST 25-APR-2001

DEFINITION 214520 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BE809138
VERSION BE809138.1 GI:10240250
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 546)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 71 row: A column: 5
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
source 1..546
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

BASE COUNT 126 a 139 c 180 g 100 t 1 others
ORIGIN

Query Match 14.9%; Score 214.6; DB 10; Length 546;
Best Local Similarity 85.9%; Pred. No. 1.4e-48;
Matches 238; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 561 GAGTCCACACCCCTGCAGGTCATCTGGGCTGTAATGCAAGAAGACACACAGTACCAG 620
|||||
Db 139 GAGTCCACACCCCTGCAGGTCATCTGGGCTGTAATGCAAGAAGACACACAGTACCAG 198
|||||

QY 621 GGCTACTGGAAGTACGGGTATGATGGCAGGACACCTTGAATTCGCTGCACACTG 680
|||||

Db 199 GGGTCTCTGGAAGTACGGGTACGATGGCAGGACCATCTTGAATTCGCGCTGAGACTG 258
|||||

QY 681 GATTGGAGACACAGAACCCAGGGCTGCCCCACCAAGCTGGAGTGGGAAGGCAACAAG 740
|||||

Db 259 GATTGGAGACACAGAGCCAGGGCCCAAGTCAACCAAGCTGGAGTGGGAAGTGAACAAG 318
|||||

QY 741 ATTCGGSCCAGGACAGACAGGGCTTACCTGGAGAGGAGCTGCCCTGCACAGCTGSCAGCAG 800
|||||

Db 319 ATTCGGSCCAGGACAGACAGGGCTTACCTGCATCGGATTCGCCCGCAGCAGCTGCTGCAT 378
|||||

QY 801 TTGCTGAGCTGGGGAGAGGTGTTTGGACCAACAAG 837
|||||

Db 379 TTGCTGAGCTGGGGAGAGGGCTCTCTGAGCAGCAAG 415
|||||

Search completed: June 19, 2002, 07:21:40
Job time: 2586 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:35 ; Search time 352.95 Seconds
(without alignments)
7004.836 Million cell updates/sec

Title: US-09-497-957-9

Perfect score:

Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAAGATTTCCTCCA 1440

Scoring table: IDENTITY NUC

scoring table.
IDENTITY_NOC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Minimum DB seq	length: 0
Maximum DB seq	length: 2000000000

Post-processing: Minimum Match 0%

Post-processing: MINIMUM MATCH 0%
Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

1:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1440	100.0	1440	18	AAT96691	Hereditary haemoch
2	1440	100.0	1440	22	AAC68429	Human hereditary h
3	1440	100.0	2727	19	AAV23525	Haemochromatosis g
4	1438.4	99.9	1440	22	AAC68430	Human hereditary h
5	1438.4	99.9	1440	22	AAC68431	Human hereditary h
6	1436.8	99.8	1440	22	AAC68432	Human hereditary h
7	1219	84.7	2506	21	AA96769	cDNA sequence enco
8	559.2	38.8	596	22	AAI63897	Human polynucleoti
9	321	22.3	10825	18	AAT96690	Hereditary haemoch

10	321	22.3	10825	22	AAC68425	Human hereditary h
11	321	22.3	10825	22	AAC68426	Human hereditary h
12	321	22.3	12146	21	AAAG6794	Genomic DNA of a h
13	321	22.3	12146	21	AAAG6794	Hereditary haemoch
14	319.4	22.2	10825	22	AAC68427	Human hereditary h
15	319.4	22.2	10825	22	AAC68428	Human hereditary h
16	280	19.4	359	20	AAAX16055	Hereditary haemoch
17	280	19.4	517	22	AAC68440	Human hereditary h
18	280	19.4	5749	22	AAL36747	Human musculoskele
19	278.4	19.3	517	22	AAC68441	Human hereditary h
20	191.8	13.3	8622	24	ABL34143	Human immune syste
21	178.2	12.4	8622	24	ABL34143	Human immune syste
22	175.2	12.2	1112	21	AAAH8668	CDNA encoding chlc
23	173.6	12.1	1230	21	AAAH8673	CDNA encoding chlc
24	170.4	11.8	1195	21	AAAH8671	CDNA encoding chlc
25	168.4	11.7	1197	21	AAAH8672	CDNA encoding chlc
26	168.4	11.7	1262	21	AAAH8691	Consensus DNA of c
27	165.8	11.6	1173	21	AAAC78071	Human cancer assoc
28	165.8	11.6	1173	22	AAAT72767	Human prostate can
29	163.6	11.4	1145	21	AAAH8667	CDNA encoding chlc
30	162	11.2	1230	21	AAAH8669	CDNA encoding chlc
31	162	11.2	1284	9	AAN80603	Probe F10 of Major
32	160.4	11.1	1230	21	AAAH8665	CDNA encoding chlc
33	160.4	11.1	1230	21	AAAH8670	CDNA encoding chlc
34	158.8	11.0	1230	21	AAAH8666	CDNA encoding chlc
35	157.2	10.9	1230	21	AAAH8664	CDNA encoding chlc
36	145.2	10.1	1554	22	AAI93004	Human polynucleoti
37	144.8	10.1	1101	12	AAH12116	HLA-C exon Cb-1.
38	143.6	10.0	1567	22	AAH98676	Human EST-derived
39	143.4	10.0	1073	22	AAH42223	Nucleotide sequenc
40	142.8	9.9	1098	22	AAAS07697	Human cDNA encodin
41	141.6	9.8	1101	12	AAQ12117	HLA-C exon Cb-2.
42	141	9.8	2034	23	AAAS90913	DNA encoding novel
43	141	9.8	2037	23	AAAS90740	DNA encoding novel
44	140	9.7	4965	16	AAH75973	pHLA-B7/beta-2 mic
45	139.4	9.7	1089	22	AAH45555	Human cancer cell

ALIGNMENTS

RESULT	1	
AAT96691		
ID	AAT96691 standard; cDNA; 1440 BP.	
XX		
AC	AAT96691;	
XX		
DT	14-APR-1998 (first entry)	
XX		
DE	Hereditary haemochromatosis gene cDNA clone.	
XX		
KW	Hereditary haemochromatosis; metal toxicity; diagnosis;	
KW	gene therapy; prenatal screening; human; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	222..1268
FT		/*tag= a
FT	mutation	408
FT		/*tag= g
FT		/note= "C to G substitution (24d2 mutation)
FT		results in His to Asp substitution"
FT		414
FT	variation	
FT		/*tag= h
FT		/note= "A to T substitution (24d7 variant)"
FT		results in Ser to Cys substitution"
FT		1066
FT	mutation	
FT		/*tag= i
FT		/note= "G to A substitution (24d1 mutation
FT		associated with HH), results in Cys to
FT		Tyr substitution"
FT		
XX		

PN W09738137-A1.
XX 16-OCT-1997.
XX 04-APR-1997; 97WO-US06254.
XX 23-MAY-1996; 96US-0652265.
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
XX (MERC-) MERCATOR GENETICS INC.
XX
XX Drayna DT, Feder JN, Gnirke A, Ruddy D, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX WPI; 1997-512743/47.
DR P-PSDB; AAW36499.
XX
XX Hereditary haemochromatosis gene and variants - useful for diagnosis
PT and treatment of hereditary haemochromatosis disease
XX
XX Disclosure; Fig 4; 115pp; English.
XX
XX This cDNA clone, designated cDNA24, is derived from human gene
CC whose mutated form is associated with hereditary haemochromatosis
CC (HH). It was obtained from a directionally cloned plasmid-based
CC cDNA library following identification of the HH locus in the HLA
CC region of chromosome 6. A single mutation (24dl) in the HH gene
CC appears responsible for the majority of HH disease. This comprises
CC a G to A substitution that is present in 86% of affected
CC chromosomes and in 4% of unaffected chromosomes. It results in a
CC Cys to Tyr substitution in the encoded protein (see AAW36499) at a
CC critical disulphide bridge important for secondary structure. The
CC following are claimed: a 10825 bp genomic DNA sequence (1) (see
CC AAW96690), the 1437 bp cDNA sequence (1a) and their 24dl, 24d2 and
CC 24d7 variants; a cloning or expression vector; host cells; a
CC peptide product chosen from the HH gene product, its variants
CC (24dl, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC residues of these; an antibody produced using the peptide; a method
CC to determine the presence or absence of the common HH gene
CC mutation; an animal model for the HH disease; metal chelation
CC agents, T-cell differentiation factors and therapeutic agents for
CC the mitigation of injury due to oxidative process in vivo or
CC therapeutic agents for activity in connection with HH disease; an
CC antisense oligonucleotide directed against a transcriptional
CC product of a nucleic acid sequence as above; and oligonucleotides
CC or pairs of oligonucleotides covering a range of nucleotides from
CC (1), (1a) or their variants, useful for detecting a polymorphism in
CC the HH gene. The invention also relates to methods for screening
CC for HH homozygotes, to HH diagnosis, prenatal screening and
CC diagnosis, and therapies of HH disease, including gene therapy,
CC protein- and antibody-based therapeutics, and small molecule
XX therapeutics.
SQ Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;

Query Match 100.0%; Score 1440; DB 18; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATACCTAGTGTTCACAGCAGGTACTCTGCTGTAGGAGAGAGAGA 60
DB 1 9999acactggatccactagtgtttccaaagcagggtaccctctctgtgtaggagagaga 60
QY 61 ACTAAAGTTCTGAAGACCTGTGCTTTTACACAGGAAGTTTACTGGGCACTCCTCTGAG 120
DB 61 actaaagtctgaagacacctgtgtttccaccaggagttttactgggcatctcctgag 120
QY 21 CCTAGGCAATAGCTAGGCTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAG 180
DB 121 ccaggcaatagctgaggtgactctgtgagccatcccccgtttcccgcccccaaaag 180

QY 181 AACCGGAGATTTAACGGGGACGTCGGCCAGAGCTGGGGAATGGCCGCCGAGCAGGC 240
DB 181 aagcggagatttaacgggggacgtgcggccagagctgggaaatgggcccgcgagcagc 240
QY 241 CGCGGCTTCTCCTCTGATGCTTTTGCAGACGCGGTCTGTCAGGGGCGCTTGTGCGCTT 300
DB 241 cggcgttctctctctgctgtctttgcagaccgcggctctgcaggggcgcttgcgcgtt 300
QY 301 CACACTTCTGCACTTACCTCTTTCATGGGTGCTCAGACAGGACCTTGTCTTCTTGT 360
DB 301 cacactctctgcactacctcttcattgggtgctcagacagacccttggtcttctcgt 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACACCGCTGTCGTGTTCTATGATCATGAGAGTCGCC 420
DB 361 ttgaagctttgggctacgtggatgaccagctgttcgtttctatgatcatgagagtcgcc 420
QY 421 GTGTGAGCCCCGAACCTCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGGTGCAGC 480
DB 421 gtgtgagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgcagc 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCTGACTTCTTGGACTATTATGG 540
DB 481 tgaagtcagagctgaagggtgggatcacatgttcaactgttgactcttgactattatgg 540
QY 541 AAAATCAACAACACAGCAGGAGTCCACACCCCTGCAGGTCTATCTGGGCTGTGAAATGC 600
DB 541 aaaaTcaacaacacagcagagtgatcccccacccctgcaggtcatctcctggctgtaaatgc 600
QY 601 AAGAAGCAACAGTACCAGGGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTGG 660
DB 601 aagaagacaacagtagccagggtacttggaagtcagggatgatggcagagaccacttg 660
QY 661 AATTCGCTCCCTGACACACTGATGGATGGAGAGCAGAGAACCCAGGGCTTGGCCCAACAGC 720
DB 661 aattctgccccgacacactggattggagagcagcagaacccaggcctggcccacaaagc 720
QY 721 TGGAGTGGGAAGGACCAAGATTTCGGGCGAGCAGACAGGGCCCTACTGTGAGAGGACT 780
DB 721 tggagtgggaaggacacaagatttcggccaggcagacagggccctacctggagaggaact 780
QY 781 GCCCTGCACACTGCAGCAGTGTGCTGGAGCTGGGGAGAGGTGTTTGGACCAACAATGC 840
DB 781 gccctgcacagtgcagcagttgctggagctggggagaggtgttttggaccacaagtgc 840
QY 841 CTCTCTTGTGTTCAAGGTGACACATCATGTGACTCTTTCAGTGACCACTCTACGGGTGTCGG 900
DB 841 ctctcttggagaagtgacacatcatgtgaccttccagtgaccactctcaggtcgtcgg 900
QY 901 CTTGAACACTACTCCCCAGAACATCACCATGAAAGTGGCTGAAGGATTAAGCAGCCAAATGG 960
DB 901 ccttgaactactacccccagaaacatcccatgaaagtggctgaaggataagcagcacaatgg 960
QY 961 ATGCCAGAGGAGTTCGAACCTTAAGACGTTATGCCCCTATGGGATGGGACCTTACCAGGGCT 1020
DB 961 atgccagggagttcgaaccttaagacgtattgcccattggggatggggatgggacctacaggct 1020
QY 1021 GGATAACCTTTGGCTGTATACCCCTGGGGAAGACAGAGATATACGTGCCAGGTGGAGCACC 1080
DB 1021 ggataaccttggctgtatccccctggggaagagcagagatatatacgtgccaggtggagcacc 1080
QY 1081 CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGGACCCCTAGTCA 1140
DB 1081 cagccctggatcagccctcatttgtatctgggagccctcaccgtctggcaccctagtc 1140
QY 1141 TTGAGGTCAATCAGTGGAAATTTGCTTTTGTGCTGTCATCTTGTTCATTTGGAATTTGTTCA 1200
DB 1141 ttgaggtcaatcagtggaattgctgttttttgcgtcacttctgttcattggaaatttctgca 1200
QY 1201 TAATATTAAAGAGAGCAGGGTTCAAGAGAGCCATGGGGGCACTACGCTTACGTTAGCTGAAC 1260
DB 1201 taatattaaggaagagcaggggttcaagagagcattggggcactacgtctttagctgaac 1260

QY 1261 GTGAGTGACAGCGAGCCTGCAGACTCACTGTGGAGAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 gtgagtgacacgcagcctgcagactcaactgtggaagagagacaaactagagactcaag 1320
QY 1321 AGGAGTGCATTATAGAGCTTTCATGTTTCAGGAGAGACTTGAACCTAAACATAGAAAT 1380
Db 1321 agggagtgaattatgagctctcatgttttcaggagagagattgaaacctaaacatagaaat 1380
QY 1381 TGCTGACGAACCTCCTTGATTTAGCCTTCTCTGTTCATTTCTCAAAAAGATTTCCCCA 1440
Db 1381 tgcctgacgaactccttgattttagccttctgttcttctcaaaaagatttcccca 1440

RESULT 2
AAC68429 standard; DNA; 1440 BP.
XX AAC68429;
AC XX
AT XX
DT 21-FEB-2001 (first entry)
XX Human hereditary hemochromatosis cDNA.
DE HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX Homo sapiens.
OS US6140305-A.
PN 31-OCT-2000.
PD XX
PE 04-APR-1997; 97US-0834497.
PF XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX (BIRA) BIO-RAD LAB INC.
PA Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
PI XX
XX WPI; 2001-006341/01.
DR XX
XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 4; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;

Query Match 100.0%; Score 1440; DB 22; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGAGTACCTTCTCTGTAGAGAGAGA 60
Db 1 ggggacactggatcactagtggttcacagcagggaggtaccttctgctgtagagagaga 60
QY 61 ACTAAAGTTCTGAAAGACCTGTGCTTTTCCACGAGGAAGTTTACTGGGCATCTCCCTGAG 120
Db 61 actaaagtctgaaagacctgtgtctttccaccaggaagttttactgggcattcctctgag 120

QY 121 CCTAGGCAATAGCTGTAGGTGACTTCTTGAGGCCATCCCCGTTCCTCCGCCGCCCAAAAG 180
Db 121 cctaggcaatagctgtaggtgacttcttgagccatccccgtttccccgcgcccccacaaag 180
QY 181 AAGCGGAGATTTAACGGGAGCGTGCGCCGACAGAGCTGGGAAATGGCCCGGAGCCAGGC 240
Db 181 aagcggagatttaacgggagcgtgcgcgcagcgtgcggcgaagtggccgcgagccaggc 240
QY 241 CGCGGCTTCTCTCTCTGATGCTTTTGCAGACCGCGGCTTCAGAGGCGCTTGTGGGCTT 300
Db 241 cgcgcttctctctctgatgcttttgcagaccgcttcctgcagcggcgcttgcgtgctt 300
QY 301 CACACTCTCTGCACACTACTCTTCATGSGGTGCTCAGAGCAGGACCTTGGCTTTCCCTGT 360
Db 301 cacactctctgcactactcttctcatgggtgctcagagcaggaccttgggttcttctctgt 360
QY 361 TTGAAGCTTTTGGGCTACGTTGATGACCCAGCTGTTCTGTTTATGATCATGAGAGTCGCC 420
Db 361 ttgaagcttttgggctacgttgatgaccagctgttcgtgttctatgatcatgagagtcgcc 420
QY 421 GTGTGAGCCCCGAACTCCATGGGTTCCTAGTAGAAATTTCAAGCCAGATGTGGTGTGAGC 480
Db 421 gtgtgagccccgaactccatgggtttccagtagaatttcaagccagatgtggtgagc 480
QY 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTGACTTCTGGACTATTATGG 540
Db 481 tgaagtcagagctgaaagggtgggatacattgttcaactgttgacttctggactattatgg 540
QY 541 AAAATCAAAACACAGCAAGGAGTCCACACCCCTGCAGGTCTCTGGGCTGTGAAATGC 600
Db 541 aaaatacaaacacagcaaggagtcccaacacctgcaggtcactcctgggtgtgaaatgc 600
QY 601 AAGAAGCAACAGTACCGAGGGCTTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
Db 601 aagaagcaacagtaccgagggtacttgcgaagtacgggtatgtggcaggaccaccttg 660
QY 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGACCAAGGCGCTTGGCCCAACCAAGC 720
Db 661 aattctgccctgacacactggattggagagcagcaacccaggggcctggcccccaagc 720
QY 721 TGGAGTGGGAAGGCACAAAGATTCGGGCCAGGACAGAGAGGCGCTACTGTGAGAGGACT 780
Db 721 tggagtggaaaggcacaagattcggccaggcagaacaggggcctacctggagaggact 780
QY 781 GCCCTGCACAGCTGCACAGTTCGTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 gccctgcacagctgcagcagttgctggagctgggagaggtgttttggaccaacaagtgc 840
QY 841 CTCCTTTGGTCAAGGTGACACATCATGTGACCTCTTCAGTCAACCACTCTACGGTGTGGG 900
Db 841 ctccctttggtgaaggtgacacatcatgtgacctcttcagtgaccactctacggtgtcggg 900
QY 901 CCTTGAATACTACTCCCGCCAGAACATCACCATTGAAGTGGCTGAAGGATAAGCAGCAATGG 960
Db 901 ccttgaactactaccgccagaacaataccatgaagtggctgaagataaagcagccaatgg 960
QY 961 ATGCCAAGAGGTTCGAACCTTAAAGACGTATTGCCAATGGGGATGGGACCTTACCAGGGCT 1020
Db 961 atgccaaaggagttcgaacctaaagacgtatttgcaccaatggggatgggaacctaccaggct 1020
QY 1021 GGATAACCTTTGGCTGTACCCCTTGGGGAAGAGCAGACATATACCTGCCAGGTGGAGCACC 1080
Db 1021 ggataaactttggctgtacccttggggaagagcagagatatacgtgccaagtggagcacc 1080
QY 1081 CAGGCTTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGCTCTGGCACCTAGTCA 1140
Db 1081 caggcttggatcagccctcatgtgatactgggagccctcaccgctctggccccctagtca 1140
QY 1141 TTGGAGTCATCAGTGAATTTGCTTTTTTGTGCTCATCTTTGTCATTTGTCATTTGTTTCA 1200
Db 1141 ttggagtcatcagtgaatttgcgtttttgttttgcgtcatctgttcatgttgaattttgtca 1200
QY 1201 TAATATTAAAGAGAGCAGGGTTTCAAGAGAGGCCATGGGGCAGTACGTCTTACTGAAC 1260

Db 1201 taatattaaaggagagcagggtttcaaggaggagccatggggcactacgtcttagctgaac 1260
QY 1261 GTGAGTGACAGCGAGCCTCGAGACTCACGTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 gtgagtgacacgcagcctcgagactcacgtg9gaaggagacaaactagagactcaag 1320
QY 1321 AGGAGTGCATTATGAGCTCTTCATGTTTCAGGAGAGAGTTCAGCTAAACATAGAAAT 1380
Db 1321 agggagtgcaattatgagctcttcattggttcaggagagagtgaaacctaaacatagaaat 1380
QY 1381 TGCTCAGCAACTCCTTGATTTAGCCTTCCTGTTTCATCTTCCTCAAAAAGATTTCCCA 1440
Db 1381 tgcctgacgaactccttgattttagctctctctgtcttcattctctcaaaaagattttcccca 1440

RESULT 3
AAV23525
ID AAV23525 standard; mRNA; 2727 BP.

XX AAV23525;
XX
DT 10-JUL-1998 (first entry)
XX
DE Haemochromatosis gene.

XX
KW Hereditary haemochromatosis; HC gene; HH identification; diagnosis;
KW autosomal recessive disorder; ss.
XX
OS Homo sapiens.

XX
PN WO9807884-A1.
XX
PD 26-FEB-1998.

XX
PF 22-AUG-1997; 97WO-AU00539.
XX
PR 03-SEP-1996; 96AU-0002083.
PR 23-AUG-1996; 96AU-0001849.

XX
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
PI Busfield F, Cullen LM, Jazwinska EC, Powell LW;
XX
DR WPI; 1998-179064/16.

XX
PT Detection of autosomal recessive disorder - particularly hereditary
PT haemochromatosis, by detecting a mutation in the HC gene
XX
PS Disclosure; Page -: 32pp; English.

XX
CC This sequence represents the haemochromatosis (HC) gene. Mutations in
CC this sequence are detected using the method of the invention. The method
CC is for identifying an individual with hereditary haemochromatosis (HH) or
CC a predisposition to develop HH or to genetically pass on HH to an
CC offspring, comprising isolating a biological sample and amplifying a
CC region of genomic DNA in the biological sample encompassing all or part
CC of the DNA between markers D6S265 and D6S276, and detecting at least one
CC homozygous or heterozygous mutation in a nucleotide within the region.
CC The method can also be used for identifying an individual with an
CC autosomal recessive disorder (ARD) or predisposition to develop and/or
CC genetically pass on an ARD to an offspring, comprises isolating a
CC biological sample from the individual and screening genomic DNA in the
CC sample for the presence of a homozygous or heterozygous mutation in a
CC gene, the normal function of which, is required to prevent progression of
CC the disorder. The method(s) can be used to identify individuals that are
CC homozygous or heterozygous (carriers) for the mutation causing the ARD.
CC Especially the method is used to diagnose HH or predisposition to HH by
CC detecting a Cys282Tyr substitution. Individuals homozygous for this
CC mutation have HH and heterozygotes are potential carriers of the
CC disease.

XX
SQ Sequence, 2727 BP; 702 A; 606 C; 660 G; 759 T; 0 other;

Query Match 100.0%; Score 1440; DB 19; Length 2727;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGACACTGGATCACCTAGTGTTCACAAAGCAGGTACCTTCTCTCTAGSAGAGAGA 60
Db 1 ggggacactggatcacctagtgtttccaaagcaggtaccttctgtctagagagagaga 60
QY 61 ACTAAGTTCGAAAGACCTCTGCTTTTTCACAGGAAAGTTTACTGGGATCTCCCTGAG 120
Db 61 actaaagtctgaaagacctgtgtctttccaccaggaagttttactgggcatctcctgag 120
QY 121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
Db 121 cctaggccaatagctgtaggggtgacttctggagccatccccgtttccccccgcccccaaaag 180
QY 181 AAGCGGAGATTTAAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCGCGAGCCAGGC 240
Db 181 aagcggagatttaacgggggacgtgcggccagagctggggaatgggcccggagccaggc 240
QY 241 CGGCGCTTCTCCTCTGATGCTTTTGGCAGACCGCGGTCTCTGCAGGGGGCTTGGTGGCTT 300
Db 241 cggcgcttctcctcctgatgcttttgagaccgcgggtcctgcaggggcgcttgcgtgctt 300
QY 301 CACACTCTCTGCACTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGGTCTTTCCCTTGT 360
Db 301 cacactctctgacactacctcttcattgggtgcctcagagcaggaccttggcttcttcctgt 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTTCTGTGTCTATGATCATGAGAGTGCCTC 420
Db 361 ttgaagctttgggtacgtggatgaccagctgttctgttctatgatacatgagatgcctc 420
QY 421 GTGTGGAGCCCGCAACTCCATGGGTTCACAGTGAATTTCAAGCCAGATGTGGGTGCGAGC 480
Db 421 gtgtggagcccgcaactccaatgggtttccagtagaaaattcaagccagatgtggtcgtcagc 480
QY 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCTACTGTGTGACTTCTGGAATATTAGG 540
Db 481 tgagtcagagctctgaaagggtgggatcacatgttcaactgttgacttctggaactattag 540
QY 541 AAAATCACACACAGCAAGAGAGTCCGACACCCCTGCAGGTCATCTCGGTGCTGAAATGC 600
Db 541 aaaatcacacacagcaaggagtccccacaccctgcgggtccatcctcgtggtgaaatgc 600
QY 601 AAGAAAGCAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG 660
Db 601 aagaagacaacagtaaccgaggggtacttgggaagtacgggtatgaggcagaccaccttg 660
QY 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGAACCCAGGGCCTGGCCCAACCAAGC 720
Db 661 aattctgacctgacacactggattggagagcagagaaccagggctggtgcccaccagaac 720
QY 721 TGGAGTGGGAAGGCACAAAGATTCGGCCAGCAGACAGGCGCTACTCTGAGAGGGACT 780
Db 721 tggagtgggaaggcacaaagatttcggccagggcagacaggccctacctggagaggact 780
QY 781 GCCCTGCACACTGCAGCAGTTCGTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db 781 gccctgcacagctgcagcagttgctgagctgggagagtggtttttggaccacaagtgc 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACACACTCTTACGGTGTGGG 900
Db 841 ctcccttggtagaagtgacacatcatgtacctcttcagtgaaccaactctacggtgtcgtg 900
QY 901 CTTTGAAGTACTACCCCGACAGACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGG 960
Db 901 ccttgaagtactacccccagaacatcacctgaagtggctggaaggaataagcagccaatgg 960
QY 961 ATGCCAAGGAGTTCGAACCTAAAGACGCTATTGCCCCAATGGGATGGGACCTTACCGAGGCT 1020
Db 961 atgccaaaggagtttcgaacacctaaagacgtattgccccaatggggatgggacctaccagggct 1020

QY 1021 GGATAACCTGGCTGTACCCCTGGGGAAGACAGAGATATACGTGCCAGGTGGAGCAC 1080
DQ 1021 ggataaaccttggctgtaccccttggggaagagagagataacgtgccaggtggagcacc 1080
QY 1081 CAGGCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCTAGTCA 1140
DQ 1081 caggcttggatcagccctcatattgtgatctgggagccctcaccgtctggcaccctagtca 1140
QY 1141 TTGAGTCATCAGTGAATTCGTCTTTTTCGTCATCTTTGTTTCATTTGGAAATTTGTCA 1200
DQ 1141 ttgagtcatacagtgaattcgtcttttctgcatcttttgcatttgcatttgaatttgttca 1200
QY 1201 TAATATTAAAGAGAGCAGCGTTTCAAGAGAGCCATGGGCACTACGCTTTAGCTGAAC 1260
DQ 1201 taatattaaagagagcagcggttcaagagagccatggggcactacgtcttagctgaac 1260
QY 1261 GTGAGTCACAGCGCCTGCAGACTACTGTGGAGGAGACAAACTAGAGACTCAAAG 1320
DQ 1261 gtgagtcacagcgccctgcagactactgtggagaggagacaaaactagagactcaag 1320
QY 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
DQ 1321 aggagtgcatattatgagctcttcatttccagagagagattgaacctaaacatagaat 1380
QY 1381 TGCCTGAGCAACTCCTTGATTTTAGCCTTCTCTGTTCAFTTCTCAAAAAGATTTCCCCA 1440
DQ 1381 tgcctgagcaactccttgattttagccttctctgttcattctcttcctcaaaaagattcccca 1440

RESULT 4

AAC68430
ID AAC68430 standard; DNA; 1440 BP.
XX
AC AAC68430;
XX
DT 21-FEB-2001 (first entry)
XX
XX Human hereditary hemochromatosis 24d1 mutation cDNA.
DE
XX HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
KW
XX Homo sapiens.
XX
XX US6140305-A.
XX
XX 31-OCT-2000.
XX
XX 04-APR-1997; 97US-0834497.
XX
XX 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
XX (BTRA) BIO-RAD LAB INC.
FA
XX
XX Thomas WJ, Drayna DT, Gnikre A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
XX WPI; 2001-006341/01.
XX
XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 4; 108pp; English.
XX
XX The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement

CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 348 A; 355 C; 406 G; 331 T; 0 other;

Query Match 99.9%; Score 1438.4; DB 22; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACCCTAGTCTTTTCAAGCAGGTACTCTTCTGCTAGGAGAGAGAGA 60
DQ 1 ggggacacttggatcacccctagtctttcaagcaggtactcttctgctagagagagaga 60
QY 61 ACTAAGTTCTGAAGACGCTGTGCTTTTTCACCAGGAAGTTTACTTGGCATCTCCTGAG 120
DQ 61 actaagttctgaagacgctgtgcttttaccaggaagtcttacttggcatctcctgag 120
QY 121 CTTAGCAATAGCTGTAGGTCGACTTCTTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
DQ 121 cttagcaatagctgtaggtcgactcttggagccatccccgtttccccgcccccaaaag 180
QY 181 AAGCGAGATTTTAAGCGGACGTCGGGCCAGAGCTGGGAAATGGGCCCGCAGCAGGC 240
DQ 181 aagcgagattttaagcggacgtcgccagagctgggaaatggggccgagccagc 240
QY 241 CGGCGCTTCTCTCTGATGCTTTTTCAGACCGCGGTCTGCAGGGGCGCTTGTGCGTT 300
DQ 241 cggcgcttctctctgactgttttgcagaccggtctgcagggcgcttgcgtgcgtt 300
QY 301 CACACTCTCTGCACTACCTCTTCATGCGGTGCTCAGACGAGGACCTTGCTTCTCTTGT 360
DQ 301 cacactctctgcactacctcttcatggtgcctcagagcagaccttggcttctcttgt 360
QY 361 TTGAAGCTTTGGCTACGTGGATGACCTGCTGCTGCTTCTATGATCATGAGAGTCGCC 420
DQ 361 ttgaagctttggctacgtggatgacctgttgcgttctctatgatcatgagagtcgcc 420
QY 421 GTGTGGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGTCAGC 480
DQ 421 gtgtggagccccgaacctccatgggtttccagtagaatttcaagccagatgtggctgcgc 480
QY 481 TGAGTCAGAGTCTGAAGGTTGGGATCACATGTTCACTGTTGCTTCTTGGACTATTATGG 540
DQ 481 tgagtcagagctctgaagggttggatcacatgttccactgttgactcttggactata 540
QY 541 AAAATCACAAACACAGCAGAGAGTCCACACCCCTGCAGGTCATCTCTGGGCTGTGAATGC 600
DQ 541 aaaatcacaaacacagcagagagttccacacccctgcaggtcactctctggctggaatgc 600
QY 601 AAGAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGACGAGCACCTTGG 660
DQ 601 aagagacaacagtaccgagggtctactggaagtacgggtatgatgggagaccaccttg 660
QY 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGAGACCCAGGCGCTTGGCCCAACAAGC 720
DQ 661 aattctgccctgacacactggattggagagcagagacccaggcccttggcccaacagc 720
QY 721 TGGAGTGGGAAAGGACAAAGATTCGGGCCAGGACAGCAACAGGCGCTTACCTGGAGAGGACT 780
DQ 721 tggagtggaaggacaaagattcgggccagagcagaaagggcctacctggagaggact 780
QY 781 GCCTGCACAGCTGCAGCAGTGTGGAGCTGGGAGAGGTGTTTGGACCAACAAAGTGC 840
DQ 781 gccctgcacagctgcagcaggttgcaggagctgggagaggtgttttggaccacagtc 840
QY 841 CTCCTTTGTTGAGGTGACACATCATGTACCTCTTCAGTGCACCTCTACTACGCTGTCCGG 900
DQ 841 ctcctttgttgaagtgacacatcatgtgacctcttcagtgacctctcacgtgtccgg 900
QY 901 CTTTGAACCTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATGAAGACCAATGG 960
DQ 901 ctttgaactactacccccagaacaatcaccatgaagtggctgaagataagcagccaatgg 960

|||||
Db 901 cctgaaactactccccagacaatcaccatgaagtggctgaaggataaagcagccaatgg 960
Qy 961 ATGCCAGGAGTTTCGACCTTAAGACGCTATTGCCCAATGGGATGGACCTACACAGGGCT 1020
Db 961 atgccaaggagttcgaaactaaagacgtattgcccaatgggattggacctaaccagggct 1020
Qy 1021 GGATAACCTTGGCTGTACCCCTCGGGAAGACGACAGAGATATACGTGCGAGTGAGCAAC 1080
Db 1021 ggaataaccttggtctgtacccccctgggaagagcagagatatatacgtgccaggtggagcaac 1080
Qy 1081 CAGCCCTGGATCAGCCCTCACTTGTGATCTGGGAGCCCTCACCGCTCGGCACCCCTAGTCA 1140
Db 1081 caggcctggatcagccctcattgtgatctggagccctcacogctcggcaccctagatca 1140
Qy 1141 TTGGAGTCATCAGTGGAAATTCCTGTTTTCGTCATCTGTTGATTTGGAATTTGTTCA 1200
Db 1141 ttggagtcatacagtggaattgctgttttttcgtcatcttctgttcattggaatttggatca 1200
Qy 1201 TAATATTAAAGGAGGAGCGGTTCAAGAGGAGCCATCGGGCACTAGCTCTTAGCTGAAC 1260
Db 1201 taatattaaggaagagcgagggttcaagagagagccatggggcactacgtcttagctgaac 1260
Qy 1261 GTGAGTGACACGACCTGCGACCTCACTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 gtgagtgacacgcagcctgcagactcactgtgggaaggagacaaaacttagagactcaag 1320
Qy 1321 AGGAGTGCAATTATGACCTCTTCATGTTTCAGGAGAGAGTTGAACCTTAAACATAGAAAT 1380
Db 1321 agggagtgcatctaagagctctcatgtttcagagagagagttgaaacctaaacatagaaat 1380
Qy 1381 TGCCCTGAGCAACTCCTTGATTAGCCCTTCTCTGTTCATTTCCCAAAAGATTTCGCCA 1440
Db 1381 tgcctgagaaactccttgatttttagccttctctgttctcatttccctcaaaaagatttcccca 1440
RESULT 6
AAC68432
ID AAC68432 standard; DNA; 1440 BP.
AC AAC68432;
XX
XX 21-FEB-2001 (first entry)
DT Human hereditary hemochromatosis 24dl/2 mutation cDNA.
XX
DE HH; hereditary hemochromatosis; chelation agent;
XX KW T-cell differentiation factor; iron overload; ss.
XX
XX Homo sapiens.
XX
XX US6140305-A.
PN
XX
PD 31-OCT-2000.
XX
XX 04-APR-1997; 97US-0834497.
PF
XX
XX 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
XX (BIRA) BIO-RAD LAB INC.
XX
XX Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
PI
XX WPI; 2001-006341/01.
DR
XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 4; 108pp; English.
PS

XX The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 1440 BP; 348 A; 354 C; 407 G; 331 T; 0 other;

Query Match 99.8%; Score 1436.8; DB 22; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGACACTGGATACCTAGTGTTCACAAGCAGAGTACCTTCTGCTGTAGGAGAGAGA 60
Db 1 9999gacactggatcaactagttttcacaaagcaggtacacctctgctgtaggagagaga 60

Qy 61 ACTAAAGTTCTGAAAGACCTGTGTCTTTTCAACAGGAAGTTTACTGGGCATCTCTCTAG 120
Db 61 actaaagtcttctgaaagacctgttgcctttttcacccaggaagttttactgggcatctctgag 120

Qy 121 CCTAGGCAATAGCTGTAGGCTGACTTCTGGAGCCATCCCGTTTCCGCCGCCCAAAAG 180
Db 121 cctaggcaatagctgtagggtgacttctctggagccaacctcgtttcccccccaaaaag 180

Qy 181 AAGCGGAGATTAAACGGGGACGTGCGGCCAGAGCTGGGAAATFGGCCCGCAGCCAGGC 240
Db 181 aagcggagatttaacggggagcgtgcggccagagctgggaaaatggccgcgagccaggc 240

Qy 241 CGCGGCTTCTCCTCTGATGCTTTTTCGACAGCCGGCTCTGCGAGGGCGCTGCTCGCTT 300
Db 241 cggcgcttctcctcctgatgcttttgcagaccgcggtctctgcaggggcgctgctgcgt 300

Qy 301 CACACTCTCTGCATCTACTCTTTCATGGTCCCTCAGAGCAGGACCTTGCTTCTTCTTGT 360
Db 301 cacactctctgcactacctctctcatgggtgctcagagcagaccttggcttctctgt 360

Qy 361 TTGAAGCTTTGGGCTAGCTGGATGACCGAGCTGTTCGTGTTCTATGATCATGAGAGTCGC 420
Db 361 ttgaagctttgggctacgtggatgaccagctgttctctgatgatgatgatgagagtcgccc 420

Qy 421 GTGTGAGAGCCCGAATCCATGGTTCAGTAGAATTTCAAGCCAGATGTGGCTGTGAATGC 480
Db 421 gtgtgagagcccgaaacctccatcggtttccagtagaatttcaagccagatgtggctgcagc 480

Qy 481 TGAGTCAGAGTCTGAAAGGTTGGGATCACATGTTCACTGTTGACTTCTTGCACTATTATGG 540
Db 481 tgaagtcagagctgaaagggtgggatacatttccactgttgacttctggactattatgg 540

Qy 541 AAATTCACACCACAGCAAGGAGTCCACACCTCGCAGGTCATCCTGGGCTGTGAATGC 600
Db 541 aaatcacaccacagcaaggagtccccacacctgcagggtccatcctcctgggctgfgaatagc 600

Qy 601 AAGAAGCAACAGTACCGAGGGCTTACTTGAAGTACGGGTATGATGGCAGGACCACTTG 660
Db 601 aagaagacaacagtacccagggctacttgaagtacgggtatgatggcagaccaccttg 660

Qy 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGAGAACCCAGGGCTTGCCCAACCAAGC 720
Db 661 aattctgcccctgacacactggattggagagcagacaacccaggcgctggcccaacagc 720

Qy 721 TGGAGTGGGAAGGACCAAGATTCGGCCAGGACAGAGGCTACTCTGAGAGGAGACT 780
Db 721 tggagtgggaaggacacaagattcggccaggcagacagggccctacctggagagggact 780

Qy 781 GCCCTGCACACTGCAGCAGTGTCTGGAGCTGGGGAGAGGTTGTTTGGACCAACAATGC 840
Db 781 gccctgcacactgcagcagtgctgagctgggagaggtgttttggaccaacaagtgc 840

Qy 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCTAGTGACCACCTCTACGGTGTGGG 900
Db 841 ctccttttggtgaaggtgacacatcatgtgacctcttctagtgaccacctctacgggtgtggg

PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 12-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-488781/53.
XX P-PSDB; AAM43591.
DR New isolated nucleic acids and polypeptides, useful for diagnosing,
DR treating and/or preventing human diseases and disorders -
XX Claim 1; SEQ ID NO 105; 664pp + Sequence Listing; English.
XX The invention relates to human polynucleotides (AAT63803-AAI64012) and
CC the encoded proteins (AAM434497-AAM43660) useful for preventing, treating
CC or ameliorating medical conditions e.g. by protein or gene therapy. The
CC genes were isolated from a range of human tissues disclosed in the
CC specification. The nucleic acids, proteins, antibodies and (ant)agonists
CC are useful in the diagnosis, treatment and prevention of: (a) cancer,
CC e.g. breast and ovarian cancer and other cancers of the adrenal gland,
CC bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
CC urogenital; (b) immune disorders e.g. Addison's disease, allergies,
CC autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 596 BP; 133 A; 157 C; 175 G; 126 T; 5 other;

Query Match 38.8%; Score 559.2; DB 22; Length 596;
Best Local Similarity 97.8%; Pred. No. 3e-151;
Matches 572; Conservative 4; Mismatches 8; Indels 1; Gaps 1;
Qy 166 CCCGCCCCAAAAGACGGAGATTTAACGGGACGTGGCGCAGCTGGGGAATGC 225
Db 13 cccgccccccaaaagagcgagatttaacgggacgtgcgcccagagctggggaatgg 72
Qy 226 GCCCGGAGCCAGCCGCGCTTCTCTCTGTGATGCTTTTGCAGACCGCGGCTCTGCAGG 285
Db 73 gcccgagagccagcgccgcttctctctctctctctctctctctctctctctctcagg 132
Qy 286 GCGCTTGTGCTGCTACACTCTGTGCTACCTCTTCATGGGTGCTCAGACGAGGACC 345
Db 133 ggcgcttgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 192
Qy 346 TTGCTCTTCTTCTTGAAGCTTTGGGCTACGTGATGATGATGATGATGATGATGATG 405
Db 193 ttggctcttctctgtttgaagctttgggctacgtggtggtggtggtggtggtggtggt 252
Qy 406 ATCATGAGTCGCGGTGGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCC 465
Db 253 atcatgagtcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 312
Qy 466 AGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525
Db 313 agatgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 371
Qy 526 TCTGACTATTATGAAAATCACAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
Db 372 tctgactattatgaaaatcacaaacacagagagagagagagagagagagagagagag 431
Qy 586 TGGGCTGTGAAATGCAAGACACAGTACCGAGGCTACTGGAAGTACGGGTATGATG 645
Db 432 ggggctgtgaaatgcaagagacacacagagagagagagagagagagagagagagag 491

QY 646 GCAGGACACCTTGAATCTGCCTGACACACTGATGGAGAGCAGAGAACCCAGGG 705
|||||
Db 492 ggcaggacaccttgaaattctccctgcacacactgattggagacagacagaccagg 551
|||||
QY 706 CTGGCCCCACCAAGCTGGAGTGGGAAAGGCACAAAGATTTCGGGCCA 750
|||||
Db 552 cctggcccacacnctggagtgaggaaaggcacaagattcgggcca 596
|||||
RESULT 9
AAT96690
ID AAT96690 standard; DNA; 10825 BP.
XX
AC AAT96690;
XX
DT 14-APR-1998 (first entry)
XX
DE Hereditary haemochromatosis gene.
XX
KW Hereditary haemochromatosis; metal toxicity; diagnosis;
KW gene therapy; prenatal screening; human; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 361..7147
FT /*tag= a
FT /*note= "contains introns"
FT intron 437..3761
FT /*tag= b
FT /*number= 1
FT intron 4026..4234
FT /*tag= c
FT /*number= 2
FT intron 4511..5605
FT /*tag= d
FT /*number= 3
FT intron 5882..6039
FT /*tag= e
FT /*number= 4
FT intron 6154..7106
FT /*tag= f
FT /*number= 5
FT mutation 3872
FT /*tag= g
FT /*note= "C to G substitution (24d2 mutation)
FT results in His to Asp substitution"
FT variation 3878
FT /*tag= h
FT /*note= "A to T substitution (24d7 variant)
FT results in Ser to Cys substitution"
FT mutation 5834
FT /*tag= i
FT /*note= "G to A substitution (24d1 mutation
FT associated with HH), results in Cys to
FT Tyr substitution"
XX WO9738137-A1.
XX
XX 16-OCT-1997.
XX
XX 04-APR-1997; 97WO-US06254.
XX PF
XX 23-MAY-1996; 96US-0652265.
XX PR
XX 04-APR-1996; 96US-0630912.
XX PR
XX 16-APR-1996; 96US-0632673.
XX PR
XX (MERC-) MERCATOR GENETICS INC.
XX PA
XX Drayna DT, Feder JN, Gnirke A, Ruddy D, Thomas WJ;
XX Tsuchihashi Z, Wolff RK;
XX
XX

DR WPI; 1997-512743/47.
DR P-PSDB; AAW36499.
XX
XX Hereditary haemochromatosis gene and variants - useful for diagnosis
PT and treatment of hereditary haemochromatosis disease
XX
XX Disclosure; Fig 3; 115pp; English.
XX
CC This genomic DNA sequence corresponds to the human gene whose
CC mutated form is associated with hereditary haemochromatosis (HH).
CC To identify this novel gene, allelic association patterns were
CC determined between known markers and the HH locus in the HLA region
CC of chromosome 6. A physical clone coverage was then generated of
CC extending from D6S265, which is a marker that is centromeric of
CC HLA-A, in a telomeric direction through D6S276, a marker at which
CC the allelic association was no longer observed. A single mutation
CC (24d1) in the HH gene appears responsible for the majority of HH
CC disease. This comprises a G to A substitution that is present in
CC 86% of affected chromosomes and in 4% of unaffected chromosomes.
CC It results in a Cys to Tyr substitution in the encoded protein (see
CC AAW36499) at a critical disulphide bridge important for secondary
CC structure. The following are claimed: the HH genomic DNA (1), a
CC 1437 bp cDNA sequence (1a) (see AAT96691) and their 24d1, 24d2 and
CC 24d7 variants; a cloning or expression vector; host cells; a
CC peptide product chosen from the HH gene product, its variants
CC (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid
CC residues of these; an antibody produced using the peptide; a method
CC to determine the presence or absence of the common HH gene
CC mutation; an animal model for the HH disease; metal chelation
CC agents, T-cell differentiation factors and therapeutic agents for
CC the mitigation of injury due to oxidative process in vivo or
CC mitigation of iron overload; a method for screening potential
CC therapeutic agents for activity in connection with HH disease; an
CC antisense oligonucleotide directed against a transcriptional
CC product of a nucleic acid sequence as above; and oligonucleotides
CC or pairs of oligonucleotides covering a range of nucleotides from
CC (1), (1a) or their variants, useful for detecting a polymorphism in
CC the HH gene. The invention also relates to methods for screening
CC for HH homozygotes, to HH diagnosis, prenatal screening and
CC diagnosis, and therapies of HH disease, including gene therapy,
CC protein- and antibody-based therapeutics, and small molecule
CC therapeutics.
XX
SQ Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;

Query Match 22.3%; Score 321; DB 18; Length 10825;
Best Local Similarity 72.1%; Pred. No. 8.6e-82;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;
QY 298 GTTCACACTCTCTGCACCTACCTCTTCATCGGTGCTCAGACAGGACCTTGGTCTTTCCT 357
|||||
Db 3762 gtccacactctctgcactacacctcttcacgtggctcagacagacacctggtcttct 3821
QY 358 TGCTTTGAAGCTTTGGCTACGTTGGATGACACAGCTGCTGCTGTTCTATGATCATGAGAGTC 417
|||||
Db 3822 tggttgaagctttgggctacgtggatgaccagctgttcgtttctatgatcatgagagtc 3881
QY 418 GCCGTGTGGAGCCCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGCGGTGC 477
|||||
Db 3882 gccgtgtgagccccgaactccatggttccagtagaatttcaagccagatgtggtcgc 3941
QY 478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCTACTGTTGACTCTGGACTATTA 537
|||||
Db 3942 agctgagtcagagctctgaaagggtggatcacatgttcactgttgactctcgtgactatta 4001
QY 538 TGGAAAAATCACAAACCACACAGCAAG----- 560
Db 4002 tggaaaatcacaaaccacacagcaagggtatgtgagagagggggccctcactctcctgaggttgt 4061
QY 561 ----- 560
Db 4062 cagagcttttcatcttttcatctatctcttgaaagaaacagctggaagtctgaggtctgtg 4121

QY 561 ----- 560
Db 4122 ggaagcagggaagggaattgttctctagatcatttgctctgggagtggtg 4181
QY 561 -----GAGTCCCA 568
Db 4182 aaatagggaacctattcttgggtgcagtttaaceaggctgggatttttccagagtc 4241
QY 569 CACCTCCAGGTACCTCGGCTGTGAATGCAAGAGACAACAGTACCGAGGCTACTG 628
Db 4242 caccctgcaggtcatctctggcgtgtaataatgcagaagacacagtagccgaggtactg 4301
QY 629 GAAGTACGGGTATGATGGCAGGACCACTTGTAATTCCTGCTGACACACTGATTTGGAG 688
Db 4302 gaagtacggtgatgagcagaccacttgattctgccttgacacactgagattggag 4361
QY 689 AGCAGCAGAACCCAGGCGCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAGATTGGGC 748
Db 4362 agcagcagaacccaggcctggccaccaggctggagtggaaggcacaagattcgggc 4421
QY 749 CAGGCAGAACAGGCGCTGACCTGGAGAGGGACTGCCCTGCACAGCTGCAGCTGCTGGA 808
Db 4422 caggcagaacaggcctaccctggagaggagactgccctgcacagctgcagcagttgctgga 4481
QY 809 GCTGGGAGAGGTGTTTGGACCAACAAG 837
Db 4482 gctgggagaggtgttttggaccaacaag 4510

RESULT 10

AAC68425

ID AAC68425 standard; DNA; 10825 BP.

XX

AC AAC68425;

XX

DT 21-FEB-2001 (first entry)

XX

DE Human hereditary hemochromatosis DNA.

XX

KW HH; hereditary hemochromatosis; chelation agent;

KW T-cell differentiation factor; iron overload; ds.

XX

OS Homo sapiens.

XX

PN US6140305-A.

XX

XX 31-OCT-2000.

PD

XX

PF 04-APR-1997; 97US-0834497.

XX

PR 04-APR-1996; 96US-0630912.

PR

PR 16-APR-1996; 96US-0632673.

PR

PR 23-MAY-1996; 96US-0652265.

XX

PA (BIRA) BIO-RAD LAB INC.

XX

PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;

PI Feder JN;

XX

DR WPI; 2001-006341/01.

DR

DR P-PSDB; AAB36869.

XX

PT New hereditary hemochromatosis gene products or polypeptides, useful

PT for treating hereditary hemochromatosis in a patient, and as a metal

PT chelation agent alleviating iron overload -

XX

PS Disclosure; Fig 3; 108pp; English.

XX

CC The present invention relates to hereditary hemochromatosis gene

CC products. These proteins may be used to treat a patient diagnosed as

CC having human hemochromatosis disease. It is also useful as a metal

CC chelation agent or as a T-cell differentiation factor, and for

CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;

Query Match 22.3%; Score 321; DB 22; Length 10825;
Best Local Similarity 72.1%; Pred. No. 8,6e-82;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;

QY 298 GTTCACACTCTCTGCACACTACCTCTTCATGGGTGCCCTCAGAGCAGACCTTGGTCTTTCT 357

Db 3762 gtccacactctctgcactacactcttcacgtgctcagcagcagacactgtcttctct 3821

QY 358 TGTTTGAAGCTTTGGGCTACGTGGATGACACAGCTCTTCGTGTTTATGATCATGAGATC 417

Db 3822 tgtttgaagctttgggtacgtggatgccacagctgttctgtctatgatcatgagagtc 3881

QY 418 GCGGTGTGGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477

Db 3882 gccgtgtggagcccccaactccatcctcagtagaattccaagccagatgtggctgc 3941

QY 478 AGCTCAGTCAGAGTCTGAAAGGGTGGGATCATGTTTCACTGTTTCACTTCTGGACTATTA 537

Db 3942 agctgagtcagagctcgaagggtgggatcacatgttccactgttgactctcggactatta 4001

QY 538 TGGAAATCACAACCCACACAGCAAG ----- 560

Db 4002 tggaaaatacaaccacacagcagggtatgtggagagggggcccaccctcctgaggtgtg 4061

QY 561 ----- 560

Db 4062 cagagcttttcatcttttcacatcttgatcatttgctcctgagatcatttgctctgggagtggtg 4121

QY 561 ----- 560

Db 4122 ggagcagggaagagggaaggaaatttgctcctgagatcatttgctctgggagtggtg 4181

QY 561 -----GAGTCCCA 568

Db 4182 aaatagggaacctattcttctgttgctagtttaacaaggctgggatttttccagagtc 4241

QY 569 CACCTCGAGGTACCTCGGCTGTGAATGCCAAGAGACAACAGTACCGAGGCTACTG 628

Db 4242 caccctgcaggtcatctcctggctgtgaatgcaagaagacacacagtagccgaggtactg 4301

QY 629 GAAGTACGGGTATGATGGCAGGACCACTTGAATTTGCCCTGCACACACTTGGATTGGAG 688

Db 4302 gaagtacggtgatgagcaggaccaccttgaatttgcctgcacacactggattggag 4361

QY 689 AGCAGCAGAACCCAGGCGCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAGATTCGGGC 748

Db 4362 agcagcagaacccaggcctggccaccaggctggagtggaaggcacaagattcgggc 4421

QY 749 CAGGCAGAACAGGCGCTACTCTGGAGAGGAGTGCCTTCACAGCTGCACAGCTGCTGGA 808

Db 4422 caggcagaacaggcctaccctggagaggagactgccctgcacagctgcagcagttgctgga 4481

QY 809 GCTGGGAGAGGTGTTTGGACCAACAAG 837

Db 4482 gctgggagaggtgttttggaccaacaag 4510

RESULT 11

AAC68426

ID AAC68426 standard; DNA; 10825 BP.

XX

AC AAC68426;

XX

DT 21-FEB-2001 (first entry)

XX

DE Human hereditary hemochromatosis 24d1 mutation DNA.

XX

KW HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
XX Homo sapiens.
XX US6140305-A.
XX
XX PD 31-OCT-2000.
XX PF 04-APR-1997; 97US-0834497.
XX PR 04-APR-1996; 96US-0630912.
XX PR 16-APR-1996; 96US-0632673.
XX PR 23-MAY-1996; 96US-0652265.
XX
XX (BIRA) BIO-RAD LAB INC.
XX PA
XX PI Thomas WJ, Drayna DF, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
XX WPI; 2001-006341/01.
DR P-PSDB; AAB36870.
XX

XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 3; 108pp; English.
XX
XX The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
XX Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;

Query Match 22.3%; Score 321; DB 22; Length 10825;
Best Local Similarity 72.1%; Pred. No. 8.6e-82;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;

QY 298 GTTCACACTCTCTGCACCTACCTCTTCATGGTGGCTCAGACGAGACCTTGGTCTTCT 357
Db gtccacactctctgcactacctctctcatgggtgcctcagacgagcaccttggtcttctct 3821

QY 358 TGTTTGAAGCTTGGGCTACGTGGATGACACAGCTGTCTGTCTATGATCATGAGAGTC 417
Db tgtttgaagctttgggctacgttggatgaccagctgttctcgttctatgatcatgagagtc 3881

QY 418 GCGGTGTGAGCCCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGCTGC 477
Db gccgtgtggagcccgaaactccatcctcagtttccagtagaatttcaagccagatgtgctgc 3941

QY 478 AGCTGAGTCAGAGCTGGAAGGGTGGGATCACATGTTCACTGTTCACTTCGGACTATTA 537
Db agctgagtcagagctgaaagggtgggacatcatgttcaactgttgaactctgactatta 4001

QY 538 TGGAAATCACACACACACAG----- 560
Db tggaaatcacacacacacagaggtgtatgagaggggggacctcacttctcctgaggttgt 4061

QY 561 ----- 560
Db cagagcttttcatcttttcatcgtcatctgaagaaacagctggaagctctgaggtctgtg 4121

QY 561 ----- 560
Db gtagcaggaagagggaaggaatttgcctcctcctgagatcatttggctcctggggatgtgg 4181

QY 561 -----GAGTCCCA 568
|||||

Db 4182 aaataggagacctattcttggttgagtttaacaagctggggtattttccagaggtccca 4241
QY 569 CACCTCGAGGTCACTCTGGGCTGTGAAATGCAAGAGACAACAGTACCAGGGCTACTG 628
Db caccctgcaggtcatcctctggctgtgaaatgcaagaagacaacagtagccgagggctactg 4301
QY 629 GAAGTACGGGTATGATGGGACAGGACCACCTTGAATTTCTGCCCTGCACACTGGATTGGAG 688
Db gaagtacgggtatgatgggacagggaccaccttgaattctgcctgacacactggattggag 4361
QY 689 AGCAGCAGAACCCAGGGCCTGCCCCACCAAGCTGGAGTGGGAAAGGCACAAAGATTCCGGC 748
Db agcagcagaacccagggccctggccccaccagctggagtgggaaaggcacaagattcgggc 4421
QY 749 CAGGCAGAACAGGGCCTACCTGGAGAGGGAGTGCCTGCACAGCTGCAGCAGTTGCTGGA 808
Db caggcagaacaggccctacctggagaggacctgccctgcacagctgcagcagttgctggga 4481
QY 809 GCTGGGGAGAGGTCTTTTGGACCAACAAG 837
Db gctggggagaggtgttttgaccacaagaag 4510

RESULT 12
AAA96794
ID AAA96794 standard; cDNA; 12146 BP.
XX AC
XX AAA96794;
XX
XX 19-FEB-2001 (first entry)
XX Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX Human; histocompatibility iron loading protein; HFE protein;
KW major histocompatibility complex; non-classical class I gene;
KW chromosome 6p; iron disorder; haemochromatosis; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH exon 1028..1324
FT /*tag= a
FT /number= 1
FT intron 1325..4651
FT /*tag= b
FT /number= 1
FT exon 4652..4915
FT /*tag= c
FT /number= 2
FT intron 4916..5124
FT /*tag= d
FT exon 5125..5400
FT /*tag= e
FT /number= 3
FT intron 5401..6493
FT /*tag= f
FT /number= 3
FT exon 6494..6769
FT /*tag= g
FT /number= 4
FT intron 6770..6927
FT /*tag= h
FT exon 6928..7041
FT /*tag= i
FT /number= 5
FT intron 7042..7994
FT /*tag= j
FT /number= 5
FT exon 7995..9050
FT /*tag= k
FT /number= 6

FT intron 9051..10205
FT /tag= 1
FT /number= 6
FT 10206..10637
FT /tag= m
FT XX
PN WO200059515-A1.
XX
XX 05-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US07982.
XX
XX 26-MAR-1999; 99US-0277457.
PR
XX (BILL-) BILLUPS-ROTHENBERG INC.
XX
XX Rothenberg BE, Sawada-Hirai R, Barton JC;
XX
XX WPI; 2000-647244/62.
DR
XX
XX Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT susceptibility to develop it, by determining the presence of a mutation
PT in exon 2 or an intron of a histocompatibility iron loading nucleic
PT acid -
XX
XX Example 1; Page 21-28; 55pp; English.
XX
XX The present sequence represents the human histocompatibility iron
CC loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC non-classical class I gene located on chromosome 6p. Mutations in the
CC gene lead to iron disorders. The specification describes a method for
CC diagnosing an iron disorder or a genetic susceptibility to develop the
CC disorder in a mammal. The method comprises determining the presence of
CC a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC is not a C to G missense mutation at nucleotide 187 of the sequence
CC given in A96769 (Genbank Accession number U60319). The presence of the
CC mutation indicates the disorder or the genetic susceptibility to the
CC disorder. The method is used to diagnose an iron disorder
CC e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX
XX Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;
SQ

Query Match 22.3%; Score 321; DB 21; Length 12146;
Best Local Similarity 72.1%; Pred. No. 9.le-82;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;
QY 298 GTTCACACTCTGCACTACCTTTCATGGTGCCCTCAGAGCAGACCTTGGTCTTTCCT 357
DB 4652 gtccacactctctgcactacctcttcctggtgctccagagcagaccttggctcttct 4711
QY 358 TGTTTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCGTTCATGATCATGAGAGTC 417
DB 4712 tgtttaagctttgggtcagtggtgaccagctgttcgtgtctatgatcatgagagtc 4771
QY 418 GCCGTGTGAGCCCCGAATCCATCGGTTTCAGTAGAATTTCAAGCCAGATGTGCGTCG 477
DB 4772 gccgtgtgagcccccgaactccatcctcaggtttccagtagaatttcaagccagatgtggctgc 4831
QY 478 AGCTGAGTCAGAGTCGTAAGGGTGGGATCAGATGTTCACTGTTGACITCTGGACTATTA 537
DB 4832 agctgagtcagagctgaaaggtgggatcacatgttctactgttgactctggactatta 4891
QY 538 TGGAAATCACACACAGCAAG----- 560
DB 4892 tggaaatcacacacacagcaaggttatgtggagagggggcctcctcctcctgaggttgt 4951
QY 561 ----- 560
DB 4952 cagagcttttcatcttttcatcatcttgaaggaaacagctggaagctctgaggtcttgtg 5011
QY 561 ----- 560

DB 5012 ggagcagggaagagggaatttgcctcctgagatcatttgcctcctgggatggtg 5071
QY 561 -----GAGTCCCA 568
DB 5072 aaatagggaacctattccttgggttgagtttaacaaggctggggtattttccagagtc 5131
QY 569 CACCTGCAGGTATCCTCGGCTGTGAAATGCAAGAGAACACAGTACCGAGGGCTACTG 628
DB 5132 caccctgcaggtcactcctcctggctgtgaaatgcaagaagacaacagtagcagggtactg 5191
QY 629 GAAGTACGGGTATGATGGCAGAGGACCCTTGAAATTCCTGCCCTCACACACTGGATTGGAG 688
DB 5192 gaagtacgggtgatggcagggacccttgaaatttgcctgcacacactggattggag 5251
QY 689 AGCAGCAACCCAGGGCCTGGCCACCAAGCTGGAGTGGGAAAGGACACAGATTCGGGC 748
DB 5252 agcagcagaacccaggccttgccccaccaggctggaatgggaaggcacaagattcgggc 5311
QY 749 CAGCAGAACAGGCGCTACCTTGAGAGGGAGTGCCTTCACAGCTGCAGCAGTGTGCTGGA 808
DB 5312 caggcagaacaggcctcacctggagagggaactgcctgcacagctgcagcagttgctgga 5371
QY 809 GCTGGGGAGAGGTCTTTTGGACCAACAAG 837
DB 5372 gctggggagaggtgttttgaccacaag 5400
RESULT 13
AAV57903/c
ID AAV57903 standard; DNA; 237326 BP.
XX
AC AAV57903;
XX
DT 21-DEC-1998 (first entry)
XX
DE Hereditary haemochromatosis subregion from an HH affected individual.
XX
KW Bovine butyrophilin; BF; human hereditary haemochromatosis; HFE;
KW diagnosis; iron metabolism; NPT3; NPT4; RORet; BTF1; BTF3;
KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW type 1 sodium transport gene; ss.
XX
OS Homo sapiens.
XX
PN WO9814466-A1.
XX
PD 09-APR-1998.
XX
PF 30-SEP-1997; 97WO-US17658.
XX
PR 07-MAY-1997; 97US-0852495.
PR 01-OCT-1996; 96US-0724394.
XX
PA (PROG-) PROGENITOR INC.
XX
PI Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
XX WPI; 1998-240014/21.
DR
XX
PT Hereditary haemochromatosis gene products - used to develop products
PT for the diagnosis and treatment of hereditary disorders in iron
PT metabolism
XX
PS Claim 1; Fig 9; 209pp; English.
XX
CC The present invention describes hereditary haemochromatosis gene
CC products from the human haemochromatosis gene. The present sequence
CC represents a hereditary haemochromatosis subregion from an hereditary
CC haemochromatosis (HH) affected individual. Also described is a
CC method to determine the presence or absence of the common hereditary
CC haemochromatosis (HFE) gene mutation in an individual comprising:
CC (a) providing DNA or RNA from the individual; and (b) assessing the

CC DNA or RNA for the presence or absence of a haplotype or genotype where
CC the presence or absence of the haplotype genotype indicates the likely
CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes Bf genes, which are homologues of the milk
CC protein butyrophilin (BT), and can be used in the production of agonists
CC and antagonists of BT function. Also described are: (1) a Roret gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) Npt3 and Npt4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
xx
SQ Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

Query Match 22.3%; Score 321; DB 19; Length 237326;
Best Local Similarity 72.1%; Pred. No. 3.9e-81;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;
QY 298 GTTCACACTCTGCGACTACCTCTTCATGGTGCCTCAGACGAGACCTTGGTCTTTTCT 357
Db 43338 GTTCACACTCTGCGACTACCTCTTCATGGTGCCTCAGACGAGACCTTGGTCTTTTCT 43279
QY 358 TGTTTGAAGCTTTGGGCTAGCTGATGACCAAGCTGTTCTGTTCTATGATCATGAGATC 417
Db 43278 TGTTTGAAGCTTTGGGCTAGCTGATGACCAAGCTGTTCTGTTCTATGATCATGAGATC 43219
QY 418 GCCGTGTGGAGCCCGCAACTCCATGGTGTCCAGTAGAATTTCAAGCCAGATGTGCTGC 477
Db 43218 GCCGTGTGGAGCCCGCAACTCCATGGTGTCCAGTAGAATTTCAAGCCAGATGTGCTGC 43159
QY 478 AGCTGAGTACAGTCTGAAAGGTTGGATCACAATGTTCTACTGTTGACTCTGGACTATTTA 537
Db 43158 AGCTGAGTACAGTCTGAAAGGTTGGATCACAATGTTCTACTGTTGACTCTGGACTATTTA 43099
QY 538 TGGAAATACACACCAAGCAAG----- 560
Db 43098 TGGAAATACACACCAAGCAAGGTTATGTGGAGAGGGGCGCTCACCTTCTGAGGTTGT 43039
QY 561 ----- 560
Db 43038 CAGAGCTTTTCATCTTTTCATGTCATCTTGAAGAAACAGCTGGAAGTCTGAGTCTGTG 42979
QY 561 ----- 560
Db 42978 GGAGCAGGAAGAGGGAAGGAATTTGCTTCTGAGATCATTTGCTCCTGGGATGGTG 42919
QY 561 -----GAGTCCCA 568
Db 42918 AAATAGGACCTATTCTTTGTTGACGTTAACAAGGCTGGGATTTTCCAGAGTCCCA 42859
QY 569 CACCTGACAGTATCTCTGGGCTGTGAAATGCAAGAGACAAACAGTACCGAGGGCTACTG 628
Db 42858 CACCTGACAGTATCTCTGGGCTGTGAAATGCAAGAGACAAACAGTACCGAGGGCTACTG 42799
QY 629 GAAGTACGGGTATCATGGGAGGACCACTTGAAATCTGCCTGCACACACTGGATTGGAG 688
Db 42798 GAAGTACGGGTATCATGGGAGGACCACTTGAAATCTGCCTGCACACACTGGATTGGAG 42739
QY 689 AGCAGCAAGACCCAGGGCTGGCCCAACCAAGCTGAGTGGGAAAGGCACAAAGATTTCGGGC 748
Db 42738 AGCAGCAAGACCCAGGGCTGGCCCAACCAAGCTGAGTGGGAAAGGCACAAAGATTTCGGGC 42679
QY 749 CAGGCAAGACCGGCTACCTGAGAGAGGACTGCCCTGCACAGCTGCAGAGTTGCTGGA 808
Db 42678 CAGGCAAGACCGGCTACCTGAGAGAGGAGTGCCTGCACAGCTGCAGAGTTGCTGGA 42619
QY 809 GCTGGGAGAGGTTTGGACCAACAG 837
Db 42618 GCTGGGAGAGGTTTGGACCAACAG 42590

RESULT 14
AAC68427
ID AAC68427 standard; DNA; 10825 BP.
XX
AC AAC68427;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d2 mutation DNA.
XX
KW HH: hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
XX
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI: 2001-006341/01.
DR P-PSDB; AAB36871.
XX
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure: Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;
Query Match 22.2%; Score 319.4; DB 22; Length 10825;
Best Local Similarity 72.0%; Pred. No. 2.5e-81;
Matches 539; Conservative 0; Mismatches 1; Indels 209; Gaps 1;
QY 298 GTTCACACTCTGCGACTACCTCTTCATGGTGCCTCAGACGAGACCTTGGTCTTTTCT 357
Db 3762 gtccacactctctgcaactctcttcattggtgctcagagcaggacctgtgcttcttct 3821
QY 358 TGTTTGAAGCTTTGGGCTAGCTGATGACCAAGCTGTTCTGTTCTATGATCATGAGATC 417
Db 3822 tgtttgaagctttgggctagctggtgacccagcttcgcttcattgattgagagctc 3881
QY 418 GCCGTGTGAGCCCCCAACTCCATGGTGTTCAGTAGAATTTCAAGCCAGATGTGCTGC 477
Db 3882 gccgtgtgagcccccgaactccattgggtttccagtagaatttcaagccagatggtgctc 3941
QY 478 AGCTGAGTACAGAGTCTGAAAGGTTGGATCACAATGTTCTACTGTTGACTTCTGGACTATTA 537
Db 3942 agctgagtcagagctcgaaggggtggatcacatgttcactgttgacttcttgactatta 4001
QY 538 TGGAAATACACACCAAGCAAG----- 560
Db 4002 tggaaatcacacccacagcagggttatgtgagagggggcctcaccttctcaggttgt 4061

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 06:38:35 ; Search time 79.08 Seconds
(without alignments)
4472.841 Million cell updates/sec

Title: US-09-497-957-9

Perfect score: 1440

Sequence: 1 GGGGACACTGGATCACCTAG.....TCTCAAAAGATTCCCA 1440

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents.NA.*
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1440	100.0	1440	3	US-08-652-265-9
2	1440	100.0	1440	3	US-08-834-497A-9
3	1440	100.0	1440	4	US-09-503-444A-9
4	1438.4	99.9	1440	3	US-08-652-265-10
5	1438.4	99.9	1440	3	US-08-652-265-11
6	1438.4	99.9	1440	3	US-08-834-497A-10
7	1438.4	99.9	1440	3	US-08-834-497A-11
8	1438.4	99.9	1440	4	US-09-503-444A-10
9	1438.4	99.9	1440	4	US-09-503-444A-11
10	1436.8	99.8	1440	3	US-08-652-265-12
11	1436.8	99.8	1440	3	US-08-834-497A-12
12	1436.8	99.8	1440	4	US-09-503-444A-12
13	1219	84.7	2506	4	US-09-277-457-1
14	321	22.3	10825	3	US-08-652-265-1
15	321	22.3	10825	3	US-08-652-265-3
16	321	22.3	10825	3	US-08-834-497A-1
17	321	22.3	10825	3	US-08-834-497A-3
18	321	22.3	10825	4	US-09-503-444A-1
19	321	22.3	10825	4	US-09-503-444A-3
20	321	22.3	12146	4	US-09-277-457-27
21	321	22.3	246240	2	US-08-724-394A-20
22	321	22.3	246240	2	US-08-724-394A-21
23	321	22.3	246240	2	US-08-724-394A-22
24	319.4	22.2	10825	3	US-08-652-265-5
25	319.4	22.2	10825	3	US-08-652-265-7
26	319.4	22.2	10825	3	US-08-834-497A-5
27	319.4	22.2	10825	3	US-08-834-497A-7

28 319.4 22.2 10825 4 US-09-503-444A-5 Sequence 5, Appli
29 319.4 22.2 10825 4 US-09-503-444A-7 Sequence 7, Appli
30 280 19.4 360 3 US-08-905-124-5 Sequence 5, Appli
31 280 19.4 517 1 US-08-632-673B-3 Sequence 3, Appli
32 280 19.4 517 3 US-08-652-265-20 Sequence 20, Appli
33 280 19.4 517 3 US-08-834-497A-20 Sequence 20, Appli
34 280 19.4 517 4 US-09-503-444A-20 Sequence 20, Appli
35 279 19.4 517 1 US-08-632-673B-13 Sequence 13, Appli
36 278.4 19.3 517 1 US-08-632-673B-4 Sequence 4, Appli
37 278.4 19.3 517 3 US-08-652-265-21 Sequence 21, Appli
38 278.4 19.3 517 3 US-08-834-497A-21 Sequence 21, Appli
39 278.4 19.3 517 4 US-09-503-444A-21 Sequence 5, Appli
40 175.2 12.2 1112 3 US-08-890-719-5 Sequence 5, Appli
41 173.6 12.1 1230 3 US-08-890-719-9 Sequence 10, Appli
42 170.4 11.8 1195 3 US-08-890-719-10 Sequence 7, Appli
43 168.4 11.7 1197 3 US-08-890-719-7 Sequence 37, Appli
44 168.4 11.7 1262 3 US-08-890-719-37 Sequence 4, Appli
45 163.6 11.4 1145 3 US-08-890-719-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-652-265-9
; Sequence 9, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolffi, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(408, "c")

```
; OTHER INFORMATION: /phenotype= "normal or wild-type"
; OTHER INFORMATION: (unaffected)
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(414, "a")
; OTHER INFORMATION: /phenotype= "normal or wild-type"
; OTHER INFORMATION: (unaffected)
; OTHER INFORMATION: /label= 24d7
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1066, "g")
; OTHER INFORMATION: /phenotype= "normal or wild-type"
; OTHER INFORMATION: (unaffected)
; OTHER INFORMATION: /label= 24d1
;
; US-08-652-265-9
```

```
Query Match 100.0%; Score 1440; DB 3; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACTAGTGTTCACAAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
DB 1 GGGGACACTGGATCACTAGTGTTCACAAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60

QY 61 ACTAAAGTCTGTAAGACCTGTGTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120
DB 61 ACTAAAGTCTGTAAGACCTGTGTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120

QY 121 CCTAGCAATAGCTGTAGGCTGACTTCTGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
DB 121 CCTAGCAATAGCTGTAGGCTGACTTCTGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180

QY 181 AAGCGAGATTTAAACGGGACGTCGCGCCAGAGCTGGGAAATGGCGCCGCGAGCCAGG 240
DB 181 AAGCGAGATTTAAACGGGACGTCGCGCCAGAGCTGGGAAATGGCGCCGCGAGCCAGG 240

QY 241 CGCGCTTCTCCTCCTGATGCTTTTTCAGACCGCGGTCTCAGAGGGGGCTTGTGCGGTT 300
DB 241 CGCGCTTCTCCTCCTGATGCTTTTTCAGACCGCGGTCTCAGAGGGGGCTTGTGCGGTT 300

QY 301 CACACTCTCTGCACTACCTCTTCATGGGTGCTTCAGAGCAGACCTTGTCTTCTCTTGT 360
DB 301 CACACTCTCTGCACTACCTCTTCATGGGTGCTTCAGAGCAGACCTTGTCTTCTCTTGT 360

QY 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGCTTCTGCTTCTATGATCATGAGTCGCC 420
DB 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGCTTCTGCTTCTATGATCATGAGTCGCC 420

QY 421 GTGTGAGCCCGGAACTCCATGGTTTCCAGTAGAATTTCAAGCCAGATGGGCTGCAGC 480
DB 421 GTGTGAGCCCGGAACTCCATGGTTTCCAGTAGAATTTCAAGCCAGATGGGCTGCAGC 480

QY 481 TGAGTCAGAGTCTGAAGGCTGGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGG 540
DB 481 TGAGTCAGAGTCTGAAGGCTGGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGG 540

QY 541 AAAATCAACACACAGCAGAGGAGTCCACACCTCGAGGTATCTCTGGGCTGTGAATGC 600
DB 541 AAAATCAACACACAGCAGAGGAGTCCACACCTCGAGGTATCTCTGGGCTGTGAATGC 600

QY 601 AAGAAGACAACAGTACCGAGGCTACTTGAAGTACGGTATGATGGGAGGACCACTTG 660
DB 601 AAGAAGACAACAGTACCGAGGCTACTTGAAGTACGGTATGATGGGAGGACCACTTG 660

QY 661 AATTCTGCCCTGACACTGATTGGAGAGCAGACAGACCCAGGCGCTGGCCCAACAGC 720
DB 661 AATTCTGCCCTGACACTGATTGGAGAGCAGACAGACCCAGGCGCTGGCCCAACAGC 720

QY 721 TGGAGTGGGAAGGACAGATTTCCGGCCAGCAGACAGGCGCTTACCTGGAGGGGACT 780
DB 721 TGGAGTGGGAAGGACAGATTTCCGGCCAGCAGACAGGCGCTTACCTGGAGGGGACT 780
```

```
QY 781 GCCCTGCACAGCTGCAGCAGTTCGCTGGAGCTGGGAGAGGTCTTTTGGACCAACAAGTGC 840
DB 781 GCCCTGCACAGCTGCAGCAGTTCGCTGGAGCTGGGAGAGGTCTTTTGGACCAACAAGTGC 840

QY 841 CTCCTTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTACCACCTCTACGGTGTGGG 900
DB 841 CTCCTTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTACCACCTCTACGGTGTGGG 900

QY 901 CTTTGAACCTACTACCCCGAGACATCACCATGAAGTGGCTGAAGGATGAAGCAGCAATGG 960
DB 901 CTTTGAACCTACTACCCCGAGACATCACCATGAAGTGGCTGAAGGATGAAGCAGCAATGG 960

QY 961 ATGCCAAGGAGTTTGAACCTAAAGACGATTATGCCCAATGGGATGGGACCTTACCAGGGCT 1020
DB 961 ATGCCAAGGAGTTTGAACCTAAAGACGATTATGCCCAATGGGATGGGACCTTACCAGGGCT 1020

QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCAC 1080
DB 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATACGTGCCAGGTGGAGCAC 1080

QY 1081 CAGGCTCGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
DB 1081 CAGGCTCGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

QY 1141 TTGGAGTCATCAGTGGAAATGCTGTTTTTGTGCTCATCTTTCATTTGGAAATTTTGTTC 1200
DB 1141 TTGGAGTCATCAGTGGAAATGCTGTTTTTGTGCTCATCTTTCATTTGGAAATTTTGTTC 1200

QY 1201 TAATATTAAAGAGAGGAGGTTTCAAGAGAGGATTCAGAGAGGATTCAGTCTTACCTGAAC 1260
DB 1201 TAATATTAAAGAGAGGAGGTTTCAAGAGAGGATTCAGAGAGGATTCAGTCTTACCTGAAC 1260

QY 1261 GTGAGTCACGCGACCTTGCAGACTCACTGTGGGAAGAGAGCAAACTAGAGACTCAAG 1320
DB 1261 GTGAGTCACGCGACCTTGCAGACTCACTGTGGGAAGAGAGCAAACTAGAGACTCAAG 1320

QY 1321 AGGAGTGCATTTATGAGCTTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
DB 1321 AGGAGTGCATTTATGAGCTTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

QY 1381 TGCCTGACGAACTCCTTGATTTAGCCCTTCTCTTTCATTTCTCAAAAAGATTTTCCCA 1440
DB 1381 TGCCTGACGAACTCCTTGATTTAGCCCTTCTCTTTCATTTCTCAAAAAGATTTTCCCA 1440

RESULT 2
US-08-834-497A-9
; Sequence 9, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/834,497A
FILING DATE: 04-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION NUMBER: 514
APPLICATION DATA: US 08/652,265
FILING DATE: 23-MAY-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA: 514
APPLICATION NUMBER: US 08/632,673
FILING DATE: 16-APR-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA: 514
APPLICATION NUMBER: US 08/630,912
FILING DATE: 04-APR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION: 514
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0056-999
TELECOMMUNICATION INFORMATION: 514
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "c")
OTHER INFORMATION: /phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(414, "a")
OTHER INFORMATION: /phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d7
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "g")
OTHER INFORMATION: /phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d1

US-08-834-497A-9

Query Match 100.0%; Score 1440; DB 3; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCAGTGTGTTTCCACAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
DB 1 GGGGACACTGGATCAGTGTGTTTCCACAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
QY 61 ACTAAAGTCTGAAAGACCTGTGCTTTTCCACAGGAAAGTTTACGTGGGCATCTCCTGAG 120
DB 61 ACTAAAGTCTGAAAGACCTGTGCTTTTCCACAGGAAAGTTTACGTGGGCATCTCCTGAG 120
QY 121 CCTAGGCAATAGCTAGGCTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
DB 121 CCTAGGCAATAGCTAGGCTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
QY 181 AAGCGGAGATTTAACGGGGACGTGGCGCCAGAGCTGGGAAATGGCGCCGAGCCAGGC 240
DB 181 AAGCGGAGATTTAACGGGGACGTGGCGCCAGAGCTGGGAAATGGCGCCGAGCCAGGC 240

Qy 1321 AGGAGTGCAATTAATGAGCTTTCATGTTTCAGGAGAGAGTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCAATTAATGAGCTTTCATGTTTCAGGAGAGAGTGAACCTAAACATAGAAAT 1380
Qy 1381 TGCGTACGACACCTTGGATTGAGCTTTCAGCTTCTGTTTCATTTCTCAAAAAGATTCCCA 1440
Db 1381 TGCGTACGACACCTTGGATTGAGCTTTCAGCTTCTGTTTCATTTCTCAAAAAGATTCCCA 1440

RESULT 3

US-09-503-444A-9
Sequence 9, Application US/09503444A

Patent No. 6228594

GENERAL INFORMATION:

APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "c")
OTHER INFORMATION: /phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d2
FEATURE:

NAME/KEY: allele
LOCATION: replace(414, "a")
OTHER INFORMATION: /phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d7
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "g")
OTHER INFORMATION: /phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d1
US-09-503-444A-9

Query Match 100.0%; Score 1440; DB 4; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGACACTGGATCACCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGGAGAGAGAGA 60
Db 1 GGGGACACTGGATCACCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGGAGAGAGAGA 60
Qy 61 ACTAAAGTTCGAAAGACCTGTGCTTTTACCAGGAAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCGAAAGACCTGTGCTTTTACCAGGAAAGTTTACTGGGCATCTCCTGAG 120
Qy 121 CCTAGGCAATAGCTGTAGGTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGTAGGTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
Qy 181 AAGCGGAGATTAAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGGAGCCAGGC 240
Db 181 AAGCGGAGATTAAACGGGGACGTGCGGCCAGAGCTGGGGAAATGGGCCCGGAGCCAGGC 240
Qy 241 CGGCGCTTCTCCTCTGATGCTTTTGACAGCCGGCTCTGCAGGGGGCTTGTGCGTT 300
Db 241 CGGCGCTTCTCCTCTGATGCTTTTGACAGCCGGCTCTGCAGGGGGCTTGTGCGTT 300
Qy 301 CACACTCTCTGCACCTACCTCTTTCATGGGTGCCTCAGAGCAGGACCTTGGCTTTTCCCTGT 360
Db 301 CACACTCTCTGCACCTACCTCTTTCATGGGTGCCTCAGAGCAGGACCTTGGCTTTTCCCTGT 360
Qy 361 TTGAAGCTTTGGGCTAGTGTGATCACCAGCTGTTCGTGTTCTATGATCATGAGTGCCT 420
Db 361 TTGAAGCTTTGGGCTAGTGTGATCACCAGCTGTTCGTGTTCTATGATCATGAGTGCCT 420
Qy 421 GTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Qy 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTTCACTGTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTTCACTGTGACTTCTGGACTATTATGG 540
Qy 541 AAAATCAACACACAGCAAGAGTCCACACCTTGCAGGTCATCTCGGCTGTGAATGC 600
Db 541 AAAATCAACACACAGCAAGAGTCCACACCTTGCAGGTCATCTCGGCTGTGAATGC 600
Qy 601 AAGAAGACAAACAGTACCGAGGGCTACTTGAAGTACGGGTATGATGGGAGGACCACTTG 660
Db 601 AAGAAGACAAACAGTACCGAGGGCTACTTGAAGTACGGGTATGATGGGAGGACCACTTG 660
Qy 661 AATTCTGCTTGCACACTGATTTGGAGAGCAGAGAACCCAGGGCTGGCCCAACCAAGC 720
Db 661 AATTCTGCTTGCACACTGATTTGGAGAGCAGAGAACCCAGGGCTGGCCCAACCAAGC 720
Qy 721 TGGAGTGGGAAAGCACAAGATTTCGGGCCAGGACAGACAGGCTTACCTGAGAGGAGACT 780
Db 721 TGGAGTGGGAAAGCACAAGATTTCGGGCCAGGACAGACAGGCTTACCTGAGAGGAGACT 780
Qy 781 GCCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGTGC 840

QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCCGG 900
QY 901 CCTTGAACCTACTACCCCGAGAACATCACCATGAAGTGGCTGAAGGATGAAGCAATGG 960
Db 901 CCTTGAACCTACTACCCCGAGAACATCACCATGAAGTGGCTGAAGGATGAAGCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTTAAGACGATATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTTAAGACGATATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
QY 1021 GGATACCTTTGGCTGTACCCCTCGGGAGAGCAGACAGATATACGTGCCAGGTGGAGACC 1080
Db 1021 GGATACCTTTGGCTGTACCCCTCGGGAGAGCAGACAGATATACGTGCCAGGTGGAGACC 1080
QY 1081 CAGGCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db 1081 CAGGCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
QY 1141 TTGGAGTCATCAGTGAATGCTGTTTTTGTGTCATCTTGTTCATTGGAATTTGTTCA 1200
Db 1141 TTGGAGTCATCAGTGAATGCTGTTTTTGTGTCATCTTGTTCATTGGAATTTGTTCA 1200
QY 1201 TAATATTAGGAGAGGAGGTTCAAGAGAGCCATGGGGACVACCTCTTAGCTGAAC 1260
Db 1201 TAATATTAGGAGAGGAGGTTCAAGAGAGCCATGGGGACVACCTCTTAGCTGAAC 1260
QY 1261 GTGAGTGACACGACCTGACACATCTGTGGGAGGAGACAAACATAGAGACTCAAAG 1320
Db 1261 GTGAGTGACACGACCTGACACATCTGTGGGAGGAGACAAACATAGAGACTCAAAG 1320
QY 1321 AGGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAT 1380
Db 1321 AGGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAT 1380
QY 1381 TGCCCTGACCAACTCTTGAATTTACCTCTCTGTTCAATTTCCCTCAAAAAGATTTCGCCA 1440
Db 1381 TGCCCTGACCAACTCTTGAATTTACCTCTCTGTTCAATTTCCCTCAAAAAGATTTCGCCA 1440

RESULT 4
US-08-652-265-10
; Sequence 10, Application US/08652265
; Patent No. 6025130

; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-May-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1066, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION: /label= 24d1
; US-08-652-265-10

Query Match 99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
QY 61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCCACGAGGAAGTTTACTTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTCCACGAGGAAGTTTACTTGGGCATCTCCTGAG 120
QY 121 CTTAGGCAATAGCTGTAGGTGACTTCTTGAGGCATCCCGGTTTCCCGCCGCCCAAAAG 180
Db 121 CTTAGGCAATAGCTGTAGGTGACTTCTTGAGGCATCCCGGTTTCCCGCCGCCCAAAAG 180
QY 181 AAGCGGAGATTAAACGGGAGCTGCGGCCAGAGCTGGGAAATGGCCCGCAGCCAGGC 240
Db 181 AAGCGGAGATTAAACGGGAGCTGCGGCCAGAGCTGGGAAATGGCCCGCAGCCAGGC 240
QY 241 CGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCGAGGGCGCTTCTGCGGT 300
Db 241 CGCGCTTCTCCTCCTGATGCTTTTGCAGACCGCGGTCCTGCGAGGGCGCTTCTGCGGT 300
QY 301 CACACTCTCTGCACCTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGGTCTTTCTTGT 360
Db 301 CACACTCTCTGCACCTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGGTCTTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCTGTTCTATGATCATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTCTGTTCTATGATCATGAGAGTCGCC 420
QY 421 GTGTGGAGCCCGAAGCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTCAGC 480
Db 421 GTGTGGAGCCCGAAGCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGTGGGATCAATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGTGGGATCAATGTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 AAATCACAACACAGCAGGAGTCCACACCTCGAGGTGTCCTGGGCTGTGAATGC 600
Db 541 AAATCACAACACAGCAGGAGTCCACACCTCGAGGTGTCCTGGGCTGTGAATGC 600
QY 601 AAGAAGACACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG 660
Db 601 AAGAAGACACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG 660

```
QY 661 AATTTCGCCCTGACACACTGSGATTGGAGAGCAGCAGAACCCAGGCCCTGGCCACCAAGC 720
Db 661 AATTTCGCCCTGACACACTGSGATTGGAGAGCAGCAGAACCCAGGCCCTGGCCACCAAGC 720
QY 721 TGGAGTGGGAAGGACACAGATTTCGGGCCAGGCAGACAGAGGCCCTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAGGACACAGATTTCGGGCCAGGCAGACAGAGGCCCTACCTGGAGAGGACT 780
QY 781 GCCTCCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGGTCTTTTGACCAACAAGTGC 840
Db 781 GCCTCCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGGTCTTTTGACCAACAAGTGC 840
QY 841 CTCCTTTGTGTGAAGTGCACATCATGTGACCTCTTCACTGACCACTACCGGTTCGGG 900
Db 841 CTCCTTTGTGTGAAGTGCACATCATGTGACCTCTTCACTGACCACTACCGGTTCGGG 900
QY 901 CCTTGAACCTACTACCCCAAGACATCACCATGAAGTGGCTGAAGATGAAGCAAGTGG 960
Db 901 CCTTGAACCTACTACCCCAAGACATCACCATGAAGTGGCTGAAGATGAAGCAAGTGG 960
QY 961 ATGCCAAGGAGTTCGAACCTAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGATATACGTGCCAGGTGGAGCACC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGATATACGTGCCAGGTGGAGCACC 1080
QY 1081 CAGGCCCTGGATCAGCCCTCATTGTGATCTGGAGCCCTCACCCTGTCGACCCCTAGTCA 1140
Db 1081 CAGGCCCTGGATCAGCCCTCATTGTGATCTGGAGCCCTCACCCTGTCGACCCCTAGTCA 1140
QY 1141 TTGGAGTCACTAGTGAATTCGTGTTTGTCTCATCTTGTTCATTTGGAATTTGTTCA 1200
Db 1141 TTGGAGTCACTAGTGAATTCGTGTTTGTCTCATCTTGTTCATTTGGAATTTGTTCA 1200
QY 1201 TAATATTAGGAGAGGAGGCTTCAAGAGAGCCATGGGACACTACGTTAGCTGAAC 1260
Db 1201 TAATATTAGGAGAGGAGGCTTCAAGAGAGCCATGGGACACTACGTTAGCTGAAC 1260
QY 1261 GTGAGTGACACGAGCTGCAGACTCAGTGTGGGAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTGACACGAGCTGCAGACTCAGTGTGGGAGGAGACAAACTAGAGACTCAAG 1320
QY 1321 AGGAGTGCATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCCTGACCAACTCCTGTGATTTAGCCTTCTCTGTTCATTTCTCAAAAAGATTTCCCCA 1440
Db 1381 TGCCTGACCAACTCCTGTGATTTAGCCTTCTCTGTTCATTTCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 5
US-08-652-265-11
; Sequence 11, Application us/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
```

```
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patencin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(408,"g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION: /label= 24d2
; US-08-652-265-11
```

```
Query Match 99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTAGGAGAGAGAGA 60
Db 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTAGGAGAGAGAGA 60
QY 61 ACTAAAGTCTGAAAGACCTGTTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTCTGAAAGACCTGTTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CCTAGGCAATAGCTGAGGTGACTTCTGGAGCCATCCCGCTTCCCGCCGCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGAGGTGACTTCTGGAGCCATCCCGCTTCCCGCCGCCCAAAAG 180
QY 181 AAGCGGAGATTAAACGGGGAGCTGCGGCCAGAGCTGGGGAATGGGCCCGCGAGCCAGGC 240
Db 181 AAGCGGAGATTAAACGGGGAGCTGCGGCCAGAGCTGGGGAATGGGCCCGCGAGCCAGGC 240
QY 241 CGGCGCTTCTCCTCCTGATGCTTTTTCAGACCCGGGCTCTGACAGGGCGGCTTGTGGGTT 300
Db 241 CGGCGCTTCTCCTCCTGATGCTTTTTCAGACCCGGGCTCTGACAGGGCGGCTTGTGGGTT 300
QY 301 CACACTCTCTGCACTACCTCTTCATGGTGCCTCAGACAGGACCTTGGTCTTCTCCTTGT 360
Db 301 CACACTCTCTGCACTACCTCTTCATGGTGCCTCAGACAGGACCTTGGTCTTCTCCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGATGACCAAGCTGTTTCGTGTTCTATGATCATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGATGACCAAGCTGTTTCGTGTTCTATGATCATGAGAGTCGCC 420
QY 421 GTGTGGAGCCCGCAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCGCAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGGTGGGATCACATGTTTCACTGTTGACTTCTGGACTATTATGG 540
```

Db 481 TGAGTCAGAGTCGAAGGGTGGGATCACATGTTCACTGTTGACCTCTGGACTATTATGG 540
QY 541 AAAATCACAAACACAGCAAGGAGTCCACACACCTCGAGTCTATCTGGGCTGTGAATGC 600
Db 541 AAAATCACAAACACAGCAAGGAGTCCACACACCTCGAGTCTATCTGGGCTGTGAATGC 600
QY 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGAGGACT 780
Db 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGAGGACT 780
QY 661 AATCTGCGCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGCGCTGGCCCAACAGC 720
Db 661 AATCTGCGCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGCGCTGGCCCAACAGC 720
QY 721 TGGAGTGGGAAGGCAACAAGATTGGGGCCAGGCAAGAGGCGCTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAGGCAACAAGATTGGGGCCAGGCAAGAGGCGCTACCTGGAGAGGACT 780
QY 781 GCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACAGTGC 840
Db 781 GCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACAGTGC 840
QY 841 CTCCTTTGGTGAAGTGCACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG 900
Db 841 CTCCTTTGGTGAAGTGCACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG 900
QY 901 CTTTGAACCTACTACCCCGGAGAACATCACCATGAAGTGGCTGAAGGATAGCAGCAATGG 960
Db 901 CTTTGAACCTACTACCCCGGAGAACATCACCATGAAGTGGCTGAAGGATAGCAGCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTAAAGAGCTATGCCCCAATGGGATGGGACCTACAGGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTAAAGAGCTATGCCCCAATGGGATGGGACCTACAGGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTCGCGAGTGGAGCACC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTCGCGAGTGGAGCACC 1080
QY 1081 CAGGCGCTGGATCAGCCCTCATTTGATCTGGAGGCGCTCACCGTCTGGCACCCTAGTCA 1140
Db 1081 CAGGCGCTGGATCAGCCCTCATTTGATCTGGAGGCGCTCACCGTCTGGCACCCTAGTCA 1140
QY 1141 TTGAGTCACTAGTGAATGCTGTTTTGTCGTCATCTGTTCAATGGAATTTGTTC 1200
Db 1141 TTGAGTCACTAGTGAATGCTGTTTTGTCGTCATCTGTTCAATGGAATTTGTTC 1200
QY 1201 TAATATTAGGAAGAGCGAGGTTCAAGAGGAGCCATGGGCACTACGCTTAGCTGAAC 1260
Db 1201 TAATATTAGGAAGAGCGAGGTTCAAGAGGAGCCATGGGCACTACGCTTAGCTGAAC 1260
QY 1261 GTGAGTGACACGCGCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
Db 1261 GTGAGTGACACGCGCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
QY 1321 AGGAGTGCAATTTATGAGCTCTCATGTTTCAGGAGAGAGTTGAACCTAAACATGAAT 1380
Db 1321 AGGAGTGCAATTTATGAGCTCTCATGTTTCAGGAGAGAGTTGAACCTAAACATGAAT 1380
QY 1381 TGCTGACGAACCTCTGATTTTATGCTTCTGTTCAATTTCTCTCAAAAAGATTTCCCA 1440
Db 1381 TGCTGACGAACCTCTGATTTTATGCTTCTGTTCAATTTCTCTCAAAAAGATTTCCCA 1440

RESULT 6
US-08-834-497A-10
; Sequence 10, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas

; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA: US/08/834,497A
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1066, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
; US-08-834-497A-10

Query Match 99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGGGACACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
Db 1 GGGGACACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
Qy 61 ACTAAAGTTCTGAAGACCTGTTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCTGAAGACCTGTTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120
Qy 121 CCTAGGCAATAGCTAGGCTGACTTCTGAGGACCATCCCGCTTTCCCGCCCCCAAG 180

Db 121 CCTAGCAATAGCTGAGGGTACCTTCGAGCCATCCCGCTTCCCGCCCCCAAG 180
QY 181 AAGCGAGATTTAAGCGGACCTGGGCGACAGAGCTGGGAAATGGCCCGGAGCAGGC 240
Db 181 AAGCGAGATTTAAGCGGACCTGGGCGACAGAGCTGGGAAATGGCCCGGAGCAGGC 240
QY 241 CGGCGCTTCTCTCTGATGCTTTTGCAGACCGCGGTCTCTGAGGGGCGCTTGTGCGTT 300
Db 241 CGGCGCTTCTCTCTGATGCTTTTGCAGACCGCGGTCTCTGAGGGGCGCTTGTGCGTT 300
QY 301 CACACTCTCTGACACTCTCTCATGGGTGCTCAGACGAGACCTTGTCTTCTTGT 360
Db 301 CACACTCTCTGACACTCTCTCATGGGTGCTCAGACGAGACCTTGTCTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCGAGCTGTTCTGTTCTATGATCATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCGAGCTGTTCTGTTCTATGATCATGAGAGTCGCC 420
QY 421 GTGTGAGCGCCGAACTCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGAGCGCCGAACTCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGCTGAAAGGTCGGATCACATGTTCACTGTTGACTCTGGAATATATGG 540
Db 481 TGAGTCAGAGCTGAAAGGTCGGATCACATGTTCACTGTTGACTCTGGAATATATGG 540
QY 541 AAAATCACACACAGAGAGTCCACACCGCTCAGGTATCTCGGCTGTGAATGC 600
Db 541 AAAATCACACACAGAGAGTCCACACCGCTCAGGTATCTCGGCTGTGAATGC 600
QY 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
Db 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
QY 661 AATTCGCTCCTCACACTGGAATTCGAGAGCAGACAGCCAGGCTGGCCCAACAGC 720
Db 661 AATTCGCTCCTCACACTGGAATTCGAGAGCAGACAGCCAGGCTGGCCCAACAGC 720
QY 721 TGGAGTGGGAAAGGCACAGATTCGGGCGCAGCAGAACAGGCGCTTACCTGGAGGGGACT 780
Db 721 TGGAGTGGGAAAGGCACAGATTCGGGCGCAGCAGAACAGGCGCTTACCTGGAGGGGACT 780
QY 781 GCGCTGCACAGCTGCAGAGTTCGTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 GCGCTGCACAGCTGCAGAGTTCGTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGTGCACATCATGTGACCTCTTCACTGACCACTCTACGGTGTGGG 900
Db 841 CTCCTTTGGTGAAGTGCACATCATGTGACCTCTTCACTGACCACTCTACGGTGTGGG 900
QY 901 CTTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGGATAAGCAGCAATGG 960
Db 901 CTTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGGATAAGCAGCAATGG 960
QY 961 ATGCCAAGAGTTCAACCTAAGACGATTTGCCCAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGAGTTCAACCTAAGACGATTTGCCCAATGGGATGGGACCTACCAAGGCT 1020
QY 1021 GGATAACCTTGGCTTACCCCTGGGAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
Db 1021 GGATAACCTTGGCTTACCCCTGGGAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
QY 1081 CAGGCTTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db 1081 CAGGCTTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
QY 1141 TTGAGTATCATAGTGAATGCTGTTTGTGCGTCATCTTGTTCATTTGGAATTTGTTCA 1200
Db 1141 TTGAGTATCATAGTGAATGCTGTTTGTGCGTCATCTTGTTCATTTGGAATTTGTTCA 1200
QY 1201 TAATATTAGGAAGAGGAGGCTTCAAGAGGACCATGGGCACTACGCTTACGTGAAC 1260

Db 1201 TAATATTAGGAAGAGGAGGCTTCAAGAGAGGAGCCATGGGGCACTAGCTTTAGCTGAAC 1260
QY 1261 GTGAGTGACAGCGAGCCTGCAGACTCACTGTGGGAAGAGACAAAACCTAGAGACTCAAAG 1320
Db 1261 GTGAGTGACAGCGAGCCTGCAGACTCACTGTGGGAAGAGACAAAACCTAGAGACTCAAAG 1320
QY 1321 AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGAGAGAGATTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGAGAGAGATTGAACCTAAACATAGAAAT 1380
QY 1381 TGCTGACGAACTCTTGTGATTTAGCCCTCTCTGTTTCATTTCTCAAAAAGATTTCCCCA 1440
Db 1381 TGCTGACGAACTCTTGTGATTTAGCCCTCTCTGTTTCATTTCTCAAAAAGATTTCCCCA 1440
RESULT 7
US-08-834-497A-11
; Sequence 11, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

```

; NAME/KEY: CDS
; LOCATION: 222...1268
; FEATURE:
;
; NAME/KEY: allele
; LOCATION: replace(408, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d2
US-08-834-497A-11

```

```
Query Match          99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY	1	GGGGACACTGGATACCTAGTGTTCACAAGCAGGTACCTTCGTGTAGGAGAGAGAGA	60
Db	1	GGGGACACTGGATACCTAGTGTTCACAAGCAGGTACCTTCGTGTAGGAGAGAGAGA	60
QY	61	ACTAAAGTCTCGAAAGACCTGTGCTTTTCACACGGAAGTTTTACTTGGCATCTCCTGAG	120
Db	61	ACTAAAGTCTCGAAAGACCTGTGCTTTTCACACGGAAGTTTTACTTGGCATCTCCTGAG	120
QY	121	CCTAGGCAATAGCTGTAGGGTGACTTCTTGAGGCCATCCCCTGTTCCCCGCCCCCCCAAAAG	180
Db	121	CCTAGGCAATAGCTGTAGGGTGACTTCTTGAGGCCATCCCCTGTTCCCCGCCCCCCCAAAAG	180
QY	181	AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGCCCGCGAGCCAGCG	240
Db	181	AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGCCCGCGAGCCAGCG	240
QY	241	CGCGCTCTCCCTCCTGATGCTTTTGCAGACCCGGTCTCGAGGGCGCTTGTCTGCGTT	300
Db	241	CGCGCTCTCCCTCCTGATGCTTTTGCAGACCCGGTCTCGAGGGCGCTTGTCTGCGTT	300
QY	301	CACACTCTGCACTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGTGCTTCTTCTTGT	360
Db	301	CACACTCTGCACTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGTGCTTCTTCTTGT	360
QY	361	TTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC	420
Db	361	TTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCGTGTTCTATGATCATGAGAGTCGCC	420
QY	421	GTGTGGAGCCCCGAATCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTSCAGC	480
Db	421	GTGTGGAGCCCCGAATCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTSCAGC	480
QY	481	TGAGTCAGAGTCTGAAAGGGTGGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGG	540
Db	481	TGAGTCAGAGTCTGAAAGGGTGGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGG	540
QY	541	AAATCACAACACACAGCAGAGGAGTCCACACCTCGCAGGTCACTCTGGCTGTGAAATGC	600
Db	541	AAATCACAACACACAGCAGAGGAGTCCACACCTCGCAGGTCACTCTGGCTGTGAAATGC	600
QY	601	AAGAAGACAACAGTACCGAGGGCTACTTGGAAAGTACGGGTATGATGGCGAGGACCACTTG	660
Db	601	AAGAAGACAACAGTACCGAGGGCTACTTGGAAAGTACGGGTATGATGGCGAGGACCACTTG	660
QY	661	AAATCTGCCCTGCACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCCCAACGC	720
Db	661	AAATCTGCCCTGCACACTGGATTGGAGAGCAGCAGAACCCAGGGCCTGGCCCCCAACGC	720
QY	721	TGGAGTGGGAAAGGCACAGATTCGGGCCAGGCAGACAGGGCCTACCTGGAGAGGGACT	780
Db	721	TGGAGTGGGAAAGGCACAGATTCGGGCCAGGCAGACAGGGCCTACCTGGAGAGGGACT	780
QY	781	GCCTGCACAGCTGCAAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCACCAAGTGC	840
Db	781	GCCTGCACAGCTGCAAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCACCAAGTGC	840
QY	841	CTCCTTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCCTCTACCGGTGTCCGG	900
Db	841	CTCCTTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCCTCTACCGGTGTCCGG	900

Db	841	CTCCTTTGGTGAAGGTGACACATCATGTGACCCTCTTCAGTGACCACACTCTCAGTGGTGTCCGG	900
Qy	901	CCTTGAACATACTACCCCCAAGAACAATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG	960
Db	901	CCTTGAACATACTACCCCCAAGAACAATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG	960
Qy	961	ATGCCAAGGAGTTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTTACCAGGGCT	1020
Db	961	ATGCCAAGGAGTTTCGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTTACCAGGGCT	1020
Qy	1021	GGATAAACCTTTGGCTGTACCCCCCTGGGGAAAGACAGAGATATACGTGCCAGTGGAGCACC	1080
Db	1021	GGATAAACCTTTGGCTGTACCCCCCTGGGGAAAGACAGAGATATACGTGCCAGTGGAGCACC	1080
Qy	1081	CAGCGCTGGATCAGCCCTCATTTGTGATCTGGAGCCCTCACCGTCTGGCACCCCTAGTCA	1140
Db	1081	CAGCGCTGGATCAGCCCTCATTTGTGATCTGGAGCCCTCACCGTCTGGCACCCCTAGTCA	1140
Qy	1141	TTGGAGTCATCAGTGGAAATGCTTTTTTGTCTGCTCATCTTGTTCATTGGAAATTTTGTTC	1200
Db	1141	TTGGAGTCATCAGTGGAAATGCTTTTTTGTCTGCTCATCTTGTTCATTGGAAATTTTGTTC	1200
Qy	1201	TAAATTAAGGAGAGCGAGGTTTCAAGAGGAGCCATGGGCGCACTACGCTCTTAGCTGAAC	1260
Db	1201	TAAATTAAGGAGAGCGAGGTTTCAAGAGGAGCCATGGGCGCACTACGCTCTTAGCTGAAC	1260
Qy	1261	GTGAGTGACACGCAGCGCTCGACACTCACCTGTGGGAAGGAGACAAAACCTAGAGACTCAAAG	1320
Db	1261	GTGAGTGACACGCAGCGCTCGACACTCACCTGTGGGAAGGAGACAAAACCTAGAGACTCAAAG	1320
Qy	1321	AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTTAACATAGAANT	1380
Db	1321	AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTTAACATAGAANT	1380
Qy	1381	TGCGTCAGCAACCTCCCTGATTTTAGCGTTCTCTGTTTCATTTCCCTCAAAAAAGATTTCCCCA	1440
Db	1381	TGCGTCAGCAACCTCCCTGATTTTAGCGTTCTCTGTTTCATTTCCCTCAAAAAAGATTTCCCCA	1440

RESULT 8

```

US-09-503-444A-10
; Sequence 10, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/652,265
; FILING DATE: 23-May-1996
; PRIOR APPLICATION DATA:
```

```

RESULT      8
US-09-503-444A-10
: Sequence 10, Application US/09503444A
: Patent No. 6228594
: GENERAL INFORMATION:
: APPLICANT: Thomas, Winston J.
: APPLICANT: Drayna, Dennis T.
: APPLICANT: Feder, John N.
: APPLICANT: Gnirke, Andreas
: APPLICANT: Ruddy, David
: APPLICANT: Tsuchihashi, Zenta
: APPLICANT: Wolff, Roger K.
: TITLE OF INVENTION: Hereditary Hemochromatosis Gene
: NUMBER OF SEQUENCES: 44
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds LLP
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: WordPerfect Version 8
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/503,444A
: FILING DATE: 14-Feb-2000
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/652,265
: FILING DATE: 23-May-1996
: PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d1
US-09-503-444A-10

Query Match 99.9%; Score 1438.4; DB 4; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATACCTAGTGTTCACAAAGCAGGTACCTCTGCTGAGGAGAGAGA 60
Db 1 GGGGACACTGGATACCTAGTGTTCACAAAGCAGGTACCTCTGCTGAGGAGAGAGA 60
QY 61 ACTAAAGTTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CTAGGCAATAGCTGAGGTGACTTCTGGAGCCATCCCGCTTCCCGCCCCCCCCAAAAG 180
Db 121 CTAGGCAATAGCTGAGGTGACTTCTGGAGCCATCCCGCTTCCCGCCCCCCCCAAAAG 180
QY 181 AAGCGGAGATTTAAGCGGACGTGGCCAGAGCTGGGGAATGGCCCGCAGCCAGGC 240
Db 181 AAGCGGAGATTTAAGCGGACGTGGCCAGAGCTGGGGAATGGCCCGCAGCCAGGC 240
QY 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTGTCAGGGGGCTTGTGCGGTT 300
Db 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTGTCAGGGGGCTTGTGCGGTT 300
QY 301 CACACTCTCTGCACTACCTCTTCATGGGTGCTTCAGAGCAGACCTTGCTCTTCTTGT 360
Db 301 CACACTCTCTGCACTACCTCTTCATGGGTGCTTCAGAGCAGACCTTGCTCTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTGCTGTTCTATGATCATGAGATCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTGCTGTTCTATGATCATGAGATCGCC 420
QY 421 GTGTGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATACATGTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 AAAATCAACACACAGCAGAGGAGTCCACACCCCTGCAGTCATCTCTGGGCTGTGAAATGC 600

Db 541 AAAATCAACACACAGCAGAGGAGTCCACACCCCTGCAGGTCACTCTGGGCTGTGAAATGC 600
QY 601 AAGAAGACACAGTACCAGAGGCTACTGGAAGTACGGGTATGATGGCAGGACACCTTG 660
Db 601 AAGAAGACACAGTACCAGAGGCTACTGGAAGTACGGGTATGATGGCAGGACACCTTG 660
QY 661 AATTCTGCCCTGCACACTGGATTGGAGAGCAGAGAACCCAGGGCTTGCCTCCACCAAGC 720
Db 661 AATTCTGCCCTGCACACTGGATTGGAGAGCAGAGAACCCAGGGCTTGCCTCCACCAAGC 720
QY 721 TGGAGTGGAAAGGCACAAGATTCGGGCCAGGACAGAACAGGGCTTACCTGGAGAGGACT 780
Db 721 TGGAGTGGAAAGGCACAAGATTCGGGCCAGGACAGAACAGGGCTTACCTGGAGAGGACT 780
QY 781 GCCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGGAGAGGTGTTTGGACACCAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGGAGAGGTGTTTGGACACCAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGCGG 900
QY 901 CTTTGAACCTACTACCCCGAGAATCATCCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
Db 901 CTTTGAACCTACTACCCCGAGAATCATCCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTAAAGACGTAATTCGCCAATGGGATGGGACCTTACCAGGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTAAAGACGTAATTCGCCAATGGGATGGGACCTTACCAGGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATAGCTGCCAGGTGGAGCAC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATAGCTGCCAGGTGGAGCAC 1080
QY 1081 CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGGACCCCTAGTCA 1140
Db 1081 CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGGACCCCTAGTCA 1140
QY 1141 TTGGAGTCATCAGTGAATGCTGCTTTTGTGCTCATCTTTGTCATTTGGAATTTGTTC 1200
Db 1141 TTGGAGTCATCAGTGAATGCTGCTTTTGTGCTCATCTTTGTCATTTGGAATTTGTTC 1200
QY 1201 TAATATTAAAGAGAGGAGGTTTCAAGAGAGCCATGGGGGACTAGCTCTTAGCTGAAC 1260
Db 1201 TAATATTAAAGAGAGGAGGTTTCAAGAGAGCCATGGGGGACTAGCTCTTAGCTGAAC 1260
QY 1261 GTGAGTCAGCGCAGCTGCAGACTCACTGTGGGAAGAGACAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTCAGCGCAGCTGCAGACTCACTGTGGGAAGAGACAAACTAGAGACTCAAG 1320
QY 1321 AGGAGTGCAATTTATGAGCTCTTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCAATTTATGAGCTCTTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCTGAGGAACCTCCTTGATTTTACGCTTCTGTTTCATTTCTCAAAAAGATTTCCTCA 1440
Db 1381 TGCTGAGGAACCTCCTTGATTTTACGCTTCTGTTTCATTTCTCAAAAAGATTTCCTCA 1440

RESULT 9
US-09-503-444A-11
; Sequence 11, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene

NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
ATTORNEY/AGENT INFORMATION:
NAME: Polissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d2
US-09-503-444A-11

Query Match 99.9%; Score 1438.4; DB 4; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGGACACTGGATCACCTAGTGTGTTTCAAGCAGGTAACCTTCTGCTAGGAGAGAGA 60
Db 1 GGGGACACTGGATCACCTAGTGTGTTTCAAGCAGGTAACCTTCTGCTAGGAGAGAGA 60

Qy 61 ACTAAAGTCTGAAGACCTGTGCTTTTACCAGGAGAGTTTACTGGGCACTCTCCTGAG 120
Db 61 ACTAAAGTCTGAAGACCTGTGCTTTTACCAGGAGAGTTTACTGGGCACTCTCCTGAG 120

Qy 121 CCTAGGCAATAGCTGTAGGTGACTTCTGGAGCCATCCCGCTTCCCGCCGCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGTAGGTGACTTCTGGAGCCATCCCGCTTCCCGCCGCCCAAAAG 180

Qy 181 AAGCGGAGATTAAACGGGAGCTGCCGCCAGAGCTGGGGAATAGGGCCCGCAGCCAGGC 240
Db 181 AAGCGGAGATTAAACGGGAGCTGCCGCCAGAGCTGGGGAATAGGGCCCGCAGCCAGGC 240

Qy 241 CGGGCTTCTCTCTGATGCTTTTGCAGACCGGGTCTGTCAGGGCGCTTGTGCTGCTT 300
Db 241 CGGGCTTCTCTCTGATGCTTTTGCAGACCGGGTCTGTCAGGGCGCTTGTGCTGCTT 300

Qy 301 CACACTCTCTGCACTACCTCTTTCATGGTGCTCAGAGCAGGACCTTGGCTTTTCCTTGT 360
Db 301 CACACTCTCTGCACTACCTCTTTCATGGTGCTCAGAGCAGGACCTTGGCTTTTCCTTGT 360

Qy 361 TTGAAGCTTTGGGCTACGTGATGACCAAGCTGTTGCTGTTCTATGATCATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGATGACCAAGCTGTTGCTGTTCTATGATCATGAGAGTCGCC 420

Qy 421 GTGTGGAGCCCGGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCGGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTTCACTGTTGACTTCTGGACTATTATGG 540

Qy 541 AAAATCAACACACAGCAGGAGTCCACACCCCTGCAGGTCTATCTGGGCTGTGAAATGC 600
Db 541 AAAATCAACACACAGCAGGAGTCCACACCCCTGCAGGTCTATCTGGGCTGTGAAATGC 600

Qy 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGAGGACT 660
Db 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGAGGACT 660

Qy 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGACAGCCAGGCGCTGGCCCAACAGC 720
Db 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGACAGCCAGGCGCTGGCCCAACAGC 720

Qy 721 TGGAGTGGGAAAGCACAAGATTGGGGCCAGGCAGACAGGCGCTTACCTGGAGAGGAGCT 780
Db 721 TGGAGTGGGAAAGCACAAGATTGGGGCCAGGCAGACAGGCGCTTACCTGGAGAGGAGCT 780

Qy 781 GCCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGAGGTGTTTGGACCAACAAGTGC 840

Qy 841 CTCCTTTGGTGAAGGTGACATCATGTGACCTCTTCACTGACCACTCTACGGTGTCCGG 900
Db 841 CTCCTTTGGTGAAGGTGACATCATGTGACCTCTTCACTGACCACTCTACGGTGTCCGG 900

Qy 901 CCTTGAACCTACTACCCCGAGACATCACATGAAGTGCCTGAAGGATAGCAGCAATGG 960
Db 901 CCTTGAACCTACTACCCCGAGACATCACATGAAGTGCCTGAAGGATAGCAGCAATGG 960

Qy 961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020

Qy 1021 GGATAACCTTGGCTGTACCCCTGGGAAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGAAAGAGCAGAGATATACGTGCCAGGTGGAGCACC 1080

Qy 1081 CAGGCTTGATCAGCCCTTCATTTGCTGCTCATCTTGTTCATTTGGAATTTTGTTC 1200
Db 1081 CAGGCTTGATCAGCCCTTCATTTGCTGCTCATCTTGTTCATTTGGAATTTTGTTC 1200

Qy 1201 TAATATTAAAGAGAGGCGAGGTTCGAAGAGGCCATGGGCGACTACGTCTTAGCTGAAC 1260
Db 1201 TAATATTAAAGAGAGGCGAGGTTCGAAGAGGCCATGGGCGACTACGTCTTAGCTGAAC 1260

Qy 1261 GTGAGTGACACCGCCTCCAGACTCAGCTGTTGGAGGAGACAAACACTAGAGACTCAAG 1320
Db 1261 GTGAGTGACACCGCCTCCAGACTCAGCTGTTGGAGGAGACAAACACTAGAGACTCAAG 1320

Qy 1321 AGGAGGTGCATTTATGAGCTTTCATGTTTTCAGGAGAGAGTGAACATAACATAGAAAT 1380

Db 1321 AGGAGTGCACTTATGAGCTTTCATGTTTCAGGAGAGAGTGAACCTAAACATAGAAAT 1380
Qy 1381 TGCCTGACGAACTCCTTGATTTAGCCTTCTCTGTCATTTCTCATAAAGATTTCCCA 1440
Db 1381 TGCCTGACGAACTCCTTGATTTAGCCTTCTCTGTCATTTCTCATAAAGATTTCCCA 1440

RESULT 10

US-08-652-265-12
; Sequence 12, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(408, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1056, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
US-08-652-265-12

Query Match 99.8%; Score 1436.8; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGACACTGGATCACCTAGTGTTCACAAAGCAGGTACCTTCTGCTAGAGAGAGAGA 60
Db 1 GGGGACACTGGATCACCTAGTGTTCACAAAGCAGGTACCTTCTGCTAGAGAGAGAGA 60
Qy 61 ACTAAAGTTCTCAAAAGACCTGTGCTTTTACACAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCTCAAAAGACCTGTGCTTTTACACAGGAAGTTTACTGGGCATCTCCTGAG 120
Qy 121 CTTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
Db 121 CTTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
Qy 181 AAGCGGAGATTTAAGCGGACGTGCGGCCAGAGCTGGGGAATGGCCGCGAGCCAGGC 240
Db 181 AAGCGGAGATTTAAGCGGACGTGCGGCCAGAGCTGGGGAATGGCCGCGAGCCAGGC 240
Qy 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTCTGAGGGCGCTTGTGCGCTT 300
Db 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTCTGAGGGCGCTTGTGCGCTT 300
Qy 301 CACACTCTCTGCACHTACCTTCTATGCGGTGCTTACAGACAGACCTTGGTCTTTCCTTGT 360
Db 301 CACACTCTCTGCACHTACCTTCTATGCGGTGCTTACAGACAGACCTTGGTCTTTCCTTGT 360
Qy 361 TTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCTGTTCTATGATCATGAGAGTCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCTGTTCTATGATCATGAGAGTCC 420
Qy 421 GTGTGAGCGCCGAACTCCATCGGTTCAGTAGAATTTCAAGCCAGATGGGTGCGAGC 480
Db 421 GTGTGAGCGCCGAACTCCATCGGTTCAGTAGAATTTCAAGCCAGATGGGTGCGAGC 480
Qy 481 TGAGTCAGAGTCTGAAAGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAAGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Qy 541 AAATCAACAACACAGCAAGGAGTCCCACACCTGCAAGTTCATCTGGGCTGTGAAATGC 600
Db 541 AAATCAACAACACAGCAAGGAGTCCCACACCTGCAAGTTCATCTGGGCTGTGAAATGC 600
Qy 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCAGGACACCTTG 660
Db 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCAGGACACCTTG 660
Qy 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAACCCAGGCGCTGGCCACCAAGC 720
Db 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAACCCAGGCGCTGGCCACCAAGC 720
Qy 721 TGGAGTGGGAAAGGCACAAGATTTCGGGCCAGGAGACAGGCGCTTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAAGGCACAAGATTTCGGGCCAGGAGACAGGCGCTTACCTGGAGAGGACT 780
Qy 781 GCCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGAGTGTGTTTGGACCAACAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGAGTGTGTTTGGACCAACAGTGC 840
Qy 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTTCTCAGTGACCACCTTACGGTGTGCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTTCTCAGTGACCACCTTACGGTGTGCGG 900
Qy 901 CCTTGAAGTACTACCCAGCAACATCAGTGAAGTGGCTGAAGGATAAGCAGCAATAGG 960
Db 901 CCTTGAAGTACTACCCAGCAACATCAGTGAAGTGGCTGAAGGATAAGCAGCAATAGG 960
Qy 961 ATGCCAAGGAGTTCGAACCTTAAAGACCTATTGCCCATGGGATGGACCTACCAGGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTTAAAGACCTATTGCCCATGGGATGGGACCTACCAGGGCT 1020
Qy 1021 GGATAACCTTTGGCTGTATACCCCTGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACC 1080
Db 1021 GGATAACCTTTGGCTGTATACCCCTGGGGAAGAGCAGAGATATACCTGCCAGGTGGAGCACC 1080

Db 601 AAGAAGACAAACAGTACCGAGGGCTACTGGAAGTACGGGTATGCGGACGAGCAACACCTTG 660
Qy 661 AATTCTGCCCTGACACACTGATTGGAGAGCAGCAGACACCGAGGCTGCGCCACCAAGC 720
Db 661 AATTCTGCCCTGACACACTGATTGGAGAGCAGCAGACACCGAGGCTGCGCCACCAAGC 720
Qy 721 TGGAGTGGGAAAGGACAAAGATTGCGGCCAGGCGACAGACAGGCGCTACTCGAGAGGACT 780
Db 721 TGGAGTGGGAAAGGACAAAGATTGCGGCCAGGCGACAGACAGGCGCTACTCGAGAGGACT 780
Qy 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
Qy 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
Qy 901 CTTTGAACCTACTACCCCGAGACATCACCATGAAGTGGCTGAAGTAAAGCAAGCAATGG 960
Db 901 CTTTGAACCTACTACCCCGAGACATCACCATGAAGTGGCTGAAGTAAAGCAAGCAATGG 960
Qy 961 ATGCCAAGGAGTTGCAACCTAAAGACCTATTGCCCAATGGGATGGGACCTACCGGCT 1020
Db 961 ATGCCAAGGAGTTGCAACCTAAAGACCTATTGCCCAATGGGATGGGACCTACCGGCT 1020
Qy 1021 GGATAACCTTGGCTGTACCCCTGGGAGAGCAGATATACGTACCGTGGAGCACC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGAGAGCAGATATACGTACCGTGGAGCACC 1080
Qy 1081 CAGGCTGGATCAGCCCTCATGTGATCTGGAGCCCTCACCCTGCGACCCCTAGTCA 1140
Db 1081 CAGGCTGGATCAGCCCTCATGTGATCTGGAGCCCTCACCCTGCGACCCCTAGTCA 1140
Qy 1141 TTGAGTCACTAGGGAATTCGTGTTTTCGTGCTCATCTTGTTCATTTGGAAATTTGTTC 1200
Db 1141 TTGAGTCACTAGGGAATTCGTGTTTTCGTGCTCATCTTGTTCATTTGGAAATTTGTTC 1200
Qy 1201 TAATATTAAGAGAGGAGGCTTCAAGAGAGCCATGGGCGACTAGCTTTAGCTGAAC 1260
Db 1201 TAATATTAAGAGAGGAGGCTTCAAGAGAGCCATGGGCGACTAGCTTTAGCTGAAC 1260
Qy 1261 GTGAGTGACAGCGCTGCAGACTACTCTGTGGAGGAGACAAAACCTAGAGACTCAAAG 1320
Db 1261 GTGAGTGACAGCGCTGCAGACTACTCTGTGGAGGAGACAAAACCTAGAGACTCAAAG 1320
Qy 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Qy 1381 TGCTGACGAACTCCTTGATTTTAGCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCA 1440
Db 1381 TGCTGACGAACTCCTTGATTTTAGCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCA 1440

RESULT 12

US-09-503-444A-12

Sequence 12, Application US/09503444A

Patent No. 6228594

GENERAL INFORMATION:

APPLICANT: Thomas, Winston J.

APPLICANT: Drayna, Dennis T.

APPLICANT: Feder, John N.

APPLICANT: Gnirke, Andreas

APPLICANT: Ruddy, David

APPLICANT: Tsuchihashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Hereditary Hemochromatosis Gene

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d1
US-09-503-444A-12

Query Match 99.8%; Score 1436.8; DB 4; Length 1440;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGGACACTGGATCATCACCTAGTGTTCACAAACAGGTACTCTTCTGCTGTAGGAGAGAGA 60
Db 1 GGGGACACTGGATCATCACCTAGTGTTCACAAACAGGTACTCTTCTGCTGTAGGAGAGAGA 60
Qy 61 ACTAAAGTTCTGAAAGACCTGTGCTTTTCCACAGGAGTTTACTGGGCATCTCTCTGAG 120
Db 61 ACTAAAGTTCTGAAAGACCTGTGCTTTTCCACAGGAGTTTACTGGGCATCTCTCTGAG 120
Qy 121 CTTAGGCAATAGCTGTAGGTTGACTTCTGGAGCCATCCCGCTTTCCCGCCCCCAAAAG 180
Db 121 CTTAGGCAATAGCTGTAGGTTGACTTCTGGAGCCATCCCGCTTTCCCGCCCCCAAAAG 180
Qy 181 AAGCGGAGATTTAACGGGAGCTGGGCGACGCTGGGCAAGTGGGAAATGGGCAATGGGCGGAGCCAGGC 240
Db 181 AAGCGGAGATTTAACGGGAGCTGGGCGACGCTGGGCAAGTGGGCAATGGGCGGAGCCAGGC 240

QY 762 GCCTACTGAGAGGACTGCCCTGCACAGCTGCAGCAGTGTGCTGGAGCTGGGAGAGT 821
Db 541 gccactcggagaggactgcccctgcacagctgcacagctgcacagctgctgagagctg 600
QY 822 GTTTTGGACCAAGTGCCTCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTG 881
Db 601 gtttggaccaacaagtgccctctttggtgaaggtgacacatcatgtgacctcttcagtg 660
QY 882 ACCACTCTAGGTTGCTGGGCTTGAACACTACTACCCGCCAGAACATCACCATGAAGTGGCTG 941
Db 661 accactctacggtgctgggcttgaaactactaccaccagaacatcaccatgaagtggctg 720
QY 942 AAGGATAAGCAGCAATGGATGTCAGAGGTTGCAACCTTAAAGACGTATTGCCCCAATGGG 1001
Db 721 aagataaagcagccaatggatgccaaggagtgccaacctaaagacgtattgcccgaatgg 780
QY 1002 GATGGGACCTACCAGGGCTGGATAAAGCTTGGCTGTACCCCTGGGGAAGAGCAGAGATAT 1061
Db 781 gatgggacctaccagggtgataaactggctgtaacccctggggaagagcagagatat 840
QY 1062 ACGTGCCAGGTGAGCAGCAGCCAGCCTGGATCAGCCCTCATTTGATCTGGGAGCCCTCA 1121
Db 841 acgtgcaggtgagcagccagccagcctggatcagccctcattgtgatctgggagccctca 900
QY 1122 CCGTCTGGCACCCTAGTCATTTGGAGTCATCAGTGGAAATGCTGTTTTGTCGTCTCTTGG 1181
Db 901 ccgtctggaacctagtcattggagtgatcatcagtggaattgctgttttgcgtcatcttg 960
QY 1182 TTCATTGGAATTTGTTCAATATTAAGGAAGAGCAGGCTTCAAGAGAGCCATGGGG 1241
Db 961 ttcattggaattttgttcataaattlaaggaagagcagcgtttcaagagagagccatgggg 1020
QY 1242 CACTACGCTTAGCTGAGTGAAGTGTGACACGACGCTGCAGACTCACTGTGGGAAGGAGA 1301
Db 1021 cactacgtcttagctgacgtgagtgacacgcagcgtcagcagcctgctgaggaagagga 1080
QY 1302 CAAACTAGAGACTCAAAGAGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGT 1361
Db 1081 caaaactagagactcaaaaggagtgatcatttatgagctcttcattgttcagagagagtg 1140
QY 1362 TGAACCTAAACATAGAAATTCCTGTGAGCACTCCTTGATTTTGGCTTCTCTCTTCATTT 1421
Db 1141 tgaacctaaacatagaaattgctgacgaactccttgatttgacctctctcttccattt 1200

RESULT 14
US-08-652-265-1
; Sequence 1, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchinashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; * COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: Join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /note= "No. 6025130mal or wild-type (unaffected)"
; OTHER INFORMATION: Hereditary Hemochromatosis (HH) gene
; OTHER INFORMATION: allele"
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) allele
; OTHER INFORMATION: CDNA (SEQ ID NO:9)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d2(C)
; OTHER INFORMATION: allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d1(G)
; OTHER INFORMATION: allele (SEQ ID NO:20)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "c")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3878, "a")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d7
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "g")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d1
; US-08-652-265-1

Query Match 22.3%; Score 321; DB 3; Length 10825;
Best Local Similarity 72.1%; Pred. No. 8.9e-83;

Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;

QY 298 GTTCACACTCTCTGCACACCTCTTCATGGGTGCTCAGACGAGACCTTGGTCTTTCCCT 357
Db 3762 GTTCACACTCTCTGCACACCTCTTCATGGGTGCTCAGACGAGACCTTGGTCTTTCCCT 3821
QY 358 TGTTCGAAGCTTTGGGCTACCTGGATGACCAAGCTTCTTCGTCTTATGATCATGAGAGTC 417
Db 3822 TGTTCGAAGCTTTGGGCTACCTGGATGACCAAGCTTCTTCGTCTTATGATCATGAGAGTC 3881
QY 418 GCGGTGTGGAGCCCCGAACCTCCATGGGTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
Db 3882 GCGGTGTGGAGCCCCGAACCTCCATGGGTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 3941
QY 478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTTCACTTCTGGACTATTA 537
Db 3942 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTTCACTTCTGGACTATTA 4001
QY 538 TGGAAATCACAAACACAGCAG-- 560
Db 4002 TGGAAATCACAAACACAGCAG-- 4061
QY 561 ----- 560
Db 4062 CAGAGCTTTTCATCTTTTCATGCATCTTGAAGGAACAGCTGGAAGCTGAGGCTTTGTG 4121
QY 561 ----- 560
Db 4122 GGAGCAGGAAGAGGGAAGGAATTTGCTCTTCATGATCATTTGGTCCCTGGGATGTGG 4181
QY 561 -----GAGTCCCA 568
Db 4182 AAATAGGACCTATTCTTTGTTGTCAGTTAAACAGGCTGGGGATTTTTCAGAGTCCCA 4241
QY 569 CACCTCTGCAGGTCATCTCTGGGCTGTGAATGCAAGAACACACAGTACCGAGGGTACTG 628
Db 4242 CACCTCTGCAGGTCATCTCTGGGCTGTGAATGCAAGAACACAGTACCGAGGGTACTG 4301
QY 629 GAAGTACGGGTATGATGGCAGGACACCTTGAATCTGCCCTGCACACAGTGGATGGAG 688
Db 4302 GAAGTACGGGTATGATGGCAGGACACCTTGAATCTGCCCTGCACACAGTGGATGGAG 4361
QY 689 AGCAGCAGAACCCAGGGCTGGCCACCAAGCTGGAGTGGGAAAGCACAAGATTGGGC 748
Db 4362 AGCAGCAGAACCCAGGGCTGGCCACCAAGCTGGAGTGGGAAAGCACAAGATTGGGC 4421
QY 749 CAGGCAGAACAGGGCTTACCTGGAGAGGAGTGCCTGTGCACAGCTGCACAGTGTGCTGA 808
Db 4422 CAGGCAGAACAGGGCTTACCTGGAGAGGAGTGCCTGTGCACAGCTGCACAGTGTGCTGA 4481
QY 809 GCTGGGGAGAGGTGTTTGGACCAACAAG 837
Db 4482 GCTGGGGAGAGGTGTTTGGACCAACAAG 4510

RESULT 15

US-08-652-265-3

; Sequence 3, Application US/08652265

; Patent No. 6025130

; GENERAL INFORMATION:

; APPLICANT: Thomas, Winston J.

; APPLICANT: Drayna, Dennis T.

; APPLICANT: Feder, John N.

; APPLICANT: Gnirke, Andreas

; APPLICANT: Ruddy, David

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Hereditary Hemochromatosis Gene

; NUMBER OF SEQUENCES: 44

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: mutation"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; OTHER INFORMATION: gene 24d1 allele"
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: 24d1 allele cDNA (SEQ ID NO:10)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
; US-08-652-265-3

Query Match

Best Local Similarity 22.3%; Score 321; DB 3; Length 10825;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;

QY 298 GTTCACACTCTCTGCACACTACCTTCTTCATGGGTGCTCAGACGAGACCTTGGTCTTTCCCT 357

Db 3762 GTTCACACTCTCTGCACACTACCTTCTTCATGGGTGCTCAGACGAGACCTTGGTCTTTCCCT 3821

QY 358 TGTTCGAAGCTTTGGGCTACCTGGATGACCAAGCTTCTTCGTCTTATGATCATGAGAGTC 417

Db 3822 TGTTCGAAGCTTTGGGCTACCTGGATGACCAAGCTTCTTCGTCTTATGATCATGAGAGTC 3881


```
QY 418 GCCGTGTGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
Db 3882 GCCGTGTGGAGCCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 3941
QY 478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTAA 537
Db 3942 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTAA 4001
QY 538 TGGAAATACAAACACACGCAAG----- 560
Db 4002 TGGAAATACAAACACACGCAAGGGGTATGTGGAGAGGGGGCCCTCACCTTCTCAGGTTGT 4061
QY 561 ----- 560
Db 4062 CAGAGCTTTTTCATCTTTTTCATGCATCTTGAAGGAACAGCTGGAAGTCTCAGGCTTTGTG 4121
QY 561 ----- 560
Db 4122 GGAGCAGGGAAGAGGGAAGGAATTTGCTTCCTGAGATCATTTTGGTCCTTGGGATGGTG 4181
QY 561 -----GAGTCCCA 568
Db 4182 AATAGGGACCTATTCTCTTTGTTGAGTTAACAGGCTGGGGATTTTCCAGAGTCCCA 4241
QY 569 CACCCTGCAGGTCTATCCTGGGCTGTGAAATGCAAGAAGACACAGTACCGAGGGCTACTG 628
Db 4242 CACCCTGCAGGTCTATCCTGGGCTGTGAAATGCAAGAAGACACAGTACCGAGGGCTACTG 4301
QY 629 GAAGTACGGGTATGATGGGCAGGACCACTTGAATTTGCCCTGACACACTGGATTGGAG 688
Db 4302 GAAGTACGGGTATGATGGGCAGGACCACTTGAATTTGCCCTGACACACTGGATTGGAG 4361
QY 689 AGCAGCAGAAACCCAGGGCCTTGCCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTGCGGC 748
Db 4362 AGCAGCAGAAACCCAGGGCCTTGCCCCACCAAGCTGGAGTGGGAAAGGCACAAGATTGCGGC 4421
QY 749 CAGGCAGAACAGGCGCTTACTTGAGAGGGNCTGCCCTGCACAGCTGCACAGTTGCTGGA 808
Db 4422 CAGGCAGAACAGGCGCTTACTTGAGAGGGNCTGCCCTGCACAGCTGCACAGTTGCTGGA 4481
QY 809 GCTGGGGAGAGTGTTTTTGGACCAACAAG 837
Db 4482 GCTGGGGAGAGTGTTTTTGGACCAACAAG 4510
```

Search completed: June 19, 2002, 08:40:03
Job time: 7288 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:23:48 ; Search time 3675.59 Seconds
(without alignments)
8198.470 Million cell updates/sec

Title: US-09-497-957-12
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCTCAAAAAGATTTCCTCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pi:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description

1	1440	100.0	1440	6	AR117796	Sequence
2	1440	100.0	1440	6	AR149466	Sequence
3	1438.4	99.9	1440	6	AR117794	Sequence
4	1438.4	99.9	1440	6	AR117795	Sequence
5	1438.4	99.9	1440	6	AR149464	Sequence
6	1438.4	99.9	1440	6	AR149465	Sequence
7	1436.8	99.8	1440	6	AR117793	Sequence
8	1436.8	99.8	1440	6	AR149463	Sequence
9	1436.8	99.8	2727	9	HSU60319	U60319 Homo sapien
10	1197.8	83.2	1280	9	HSA249335	AJ249335 Homo sapi
11	1185.8	82.3	1200	9	AF115265	AF115265 Homo sapi
12	899.8	68.7	1045	9	AF079407	AF079407 Homo sapi
13	893.8	62.1	1320	4	AY007541	AY007541 Ceratotheri
14	893.8	62.1	1320	4	AY007543	AY007543 Rhinocero
15	893.6	62.1	1319	4	AY007544	AY007544 Rhinocero
16	890.6	61.8	2332	4	AY007542	AY007542 Dicerob
17	847.8	58.9	1085	9	HSA249336	AJ249336 Homo sapi
18	783.8	54.4	1073	9	HSA249337	AJ249337 Homo sapi
19	772	53.6	1885	9	AF144242	AF144242 Homo sapi
20	709.8	49.3	823	9	AF079408	AF079408 Homo sapi
21	634.4	44.1	1706	10	RNHERHAEM	AJ001517 Rattus no
22	619.6	43.0	1529	10	MMU66849	U66849 Mus musculu
23	615.8	42.8	781	9	AF079409	AF079409 Homo sapi
24	571.4	39.7	809	9	HSA250635	AJ250635 Homo sapi
25	546.6	38.0	987	9	AF150664	AF150664 Homo sapi
26	486.8	33.8	686	9	AF144240	AF144240 Homo sapi
27	475	33.0	789	10	AF008587	AF008587 Rattus no
28	461.6	32.1	804	9	AF149804	AF149804 Homo sapi
29	447.2	31.1	819	10	AF176534	AF176534 Rattus no
30	434.8	30.2	735	9	AF115264	AF115264 Homo sapi
31	391.4	27.2	580	9	AF144243	AF144243 Homo sapi
32	321	22.3	10825	6	AR117791	AR117791 Sequence
33	321	22.3	10825	6	AR117792	AR117792 Sequence
34	321	22.3	10825	6	AR149461	AR149461 Sequence
35	321	22.3	10825	6	AR149462	AR149462 Sequence
36	320.4	22.2	492	9	AF144241	AF144241 Homo sapi
37	320.4	22.2	874	9	HSUHAH2	Y09800 H.sapiens H
38	319.4	22.2	10825	6	AR117789	AR117789 Sequence
39	319.4	22.2	10825	6	AR117790	AR117790 Sequence
40	319.4	22.2	10825	6	AR149459	AR149459 Sequence
41	319.4	22.2	10825	6	AR149460	AR149460 Sequence
42	319.4	22.2	12146	9	HSHE	Z92910 Homo sapien
43	319.4	22.2	193752	2	AL359892	AL359892 Homo sapi
44	319.4	22.2	246240	6	AR036572	AR036572 Sequence
45	319.4	22.2	246240	6	AR036573	AR036573 Sequence

ALIGNMENTS

RESULT 1

AR117796

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

AR117796	Sequence	12 from patent US 6140305.	1440 bp	DNA	linear	PAT 16-MAY-2001
AR117796	Sequence	GI:14098702				
Unknown.	Unknown.					
Unclassified.	Unclassified.					
1 (bases 1 to 1440)	1 (bases 1 to 1440)					
Thomas, W.J., Drayna, D.T., Feder, J.N., Gairke, A., Ruddy, D.,	Thomas, W.J., Drayna, D.T., Feder, J.N., Gairke, A., Ruddy, D.,					
Tsushihashi, Z. and Wolff, R.K.	Tsushihashi, Z. and Wolff, R.K.					
Hereditary hemochromatosis gene products	Hereditary hemochromatosis gene products					
Patent: US 6140305-A 12 31-OCT-2000;	Patent: US 6140305-A 12 31-OCT-2000;					
Location/Qualifiers	Location/Qualifiers					
1..1440	1..1440					
/organism="unknown"	/organism="unknown"					
348 a	354 c	407 g	331 t			
Query Match	100.0%	Score 1440;	DB 6;	Length 1440;		

Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GGGACACTGGATCACCTAGTGTGTTTCAACAGCAGGTACCTTCTCTCTCTAGGAGAGAGA	60						
Db	1	GGGACACTGGATCACCTAGTGTGTTTCAACAGCAGGTACCTTCTCTCTCTAGGAGAGAGA	60						
QY	61	ACTAAAGTTCTGAAGACCTGTGCTTTTTCACAGAGAGTTTACTGGGCATCTCCTGAG	120						
Db	61	ACTAAAGTTCTGAAGACCTGTGCTTTTTCACAGAGAGTTTACTGGGCATCTCCTGAG	120						
QY	121	CCTAGGCAATAGCTGTAGGGTGACTCTGGAGCCTACCCGTTTCCCGCCGCCCAAAAG	180						
Db	121	CCTAGGCAATAGCTGTAGGGTGACTCTGGAGCCTACCCGTTTCCCGCCGCCCAAAAG	180						
QY	181	AAGCGGAGATTAAACGGGGAGCTGCGGCCAGAGCTGGGGAAATGGGCGCCGAGCAGGC	240						
Db	181	AAGCGGAGATTAAACGGGGAGCTGCGGCCAGAGCTGGGGAAATGGGCGCCGAGCAGGC	240						
QY	241	CGGCGCTTCTCCT	300						
Db	241	CGGCGCTTCTCCT	300						
QY	301	CACACTCTCTGCACTACCTCTTTCATGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTGT	360						
Db	301	CACACTCTCTGCACTACCTCTTTCATGGTGCCTCAGAGCAGGACCTTGGTCTTTCCTGT	360						
QY	361	TTGAAGCTTTGGGCTAGCTGGATACACAGCTGTTCGTGTTTCTATGATGATGAGATCGCC	420						
Db	361	TTGAAGCTTTGGGCTAGCTGGATACACAGCTGTTCGTGTTTCTATGATGATGAGATCGCC	420						
QY	421	GTGTGGAGCCCCGAACCTCCTGTTTCCAGTGAATTTCAACCCAGATGTGCTCGAGC	480						
Db	421	GTGTGGAGCCCCGAACCTCCTGTTTCCAGTGAATTTCAACCCAGATGTGCTCGAGC	480						
QY	481	TGAGTCAGAGTCTGAAAGGGTGGATCACATGTTCACTGTGTGACTTCTGGACTATTATGG	540						
Db	481	TGAGTCAGAGTCTGAAAGGGTGGATCACATGTTCACTGTGTGACTTCTGGACTATTATGG	540						
QY	541	AAATCACACACACAGGAGTCCACACCTCGAGGTATCTCTGGGCTGTGAAATGC	600						
Db	541	AAATCACACACACAGGAGTCCACACCTCGAGGTATCTCTGGGCTGTGAAATGC	600						
QY	601	AAGAAGCAACACTACCGAGGGTACTTGAAGTACGGGTATGATGGGAGGACCACTTG	660						
Db	601	AAGAAGCAACACTACCGAGGGTACTTGAAGTACGGGTATGATGGGAGGACCACTTG	660						
QY	661	AATTCGCCCCGACACACTGGATTGGAGAGCAGACAGCCAGGGCCTGGCCCAACAGC	720						
Db	661	AATTCGCCCCGACACACTGGATTGGAGAGCAGACAGCCAGGGCCTGGCCCAACAGC	720						
QY	721	TGAGTGGGAAAGGCACAGATTCGGGCCAGGCAGAACAGGGCTACCTGGAGAGGACT	780						
Db	721	TGAGTGGGAAAGGCACAGATTCGGGCCAGGCAGAACAGGGCTACCTGGAGAGGACT	780						
QY	781	GCCTCTGCACAGCTGCACAGTGTCTGGAGCTGGGAGAGTGTGTTGGACCAACAGTGC	840						
Db	781	GCCTCTGCACAGCTGCACAGTGTCTGGAGCTGGGAGAGTGTGTTGGACCAACAGTGC	840						
QY	841	CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTCTACGGTGTGCGG	900						
Db	841	CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTCTACGGTGTGCGG	900						
QY	901	CCPTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGATTAAGCAATGG	960						
Db	901	CCPTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGATTAAGCAATGG	960						
QY	961	ATGCCAAGGAGTTGCAACCTAAAGAGCTATTGCCCAATGGGATGGGACCTACAGGGCT	1020						
Db	961	ATGCCAAGGAGTTGCAACCTAAAGAGCTATTGCCCAATGGGATGGGACCTACAGGGCT	1020						
QY	1021	GGATAACCTTGGCTGTACCCCTGGGAAAGACAGAGATATACGTACCAAGTGGAGCACC	1080						
Db	1021	GGATAACCTTGGCTGTACCCCTGGGAAAGACAGAGATATACGTACCAAGTGGAGCACC	1080						

Db	1021	GGATAACCTTTGGCTGTACCCCTGGGAAAGACAGAGATATACGTACCAAGTGGAGCACC	1080																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
----	------	---	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

QY 301 CACACTCTCTGCACCTACCTCTTCATGGTGCCTCAGACGAGGACCTTGGTCTTTCCTTGT 360
Db 301 CACACTCTCTGCACCTACCTCTTCATGGTGCCTCAGACGAGGACCTTGGTCTTTCCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTCGTGTTCTATGATGATGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTCGTGTTCTATGATGATGAGTCGCC 420
QY 421 GTGTGGAGCCCCGAACCTCCATGGTTCACAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCCGAACCTCCATGGTTCACAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGCTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGCTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 AAAATCACACACACGAGGAGTCCACACCTGCAGCTCATCTGGGCTGTGAATGC 600
Db 541 AAAATCACACACACGAGGAGTCCACACCTGCAGCTCATCTGGGCTGTGAATGC 600
QY 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACTTG 660
Db 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACTTG 660
QY 661 AATTCTGCCCTGACACACTGGATGGAGAGCAGCAACCCAGGGCTGGCCACCAGC 720
Db 661 AATTCTGCCCTGACACACTGGATGGAGAGCAGCAACCCAGGGCTGGCCACCAGC 720
QY 721 TGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGACAGGGCTACCTGGAGAGGACT 780
Db 721 TGAGTGGGAAAGGCACAAGATTCGGGCCAGGCAGACAGGGCTACCTGGAGAGGACT 780
QY 781 GCCCTGCACAGCTGCAGCAGTTCTGGAGCTGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTTCTGGAGCTGGGAGAGGTGTTTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGG 900
QY 901 CCTTGAACCTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATGAAGCACCAGTGG 960
Db 901 CCTTGAACCTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATGAAGCACCAGTGG 960
QY 961 ATGCCAAGGAGTTCGAACCTTAAGACGTATGCCCAATGGGGATGGGACCTACCGGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTTAAGACGTATGCCCAATGGGGATGGGACCTACCGGGCT 1020
QY 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGCAGAGATATACGTACCGGTGGAGCACC 1080
Db 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGCAGAGATATACGTACCGGTGGAGCACC 1080
QY 1081 CAGGCCCTGGATCAGCCCTCATTTGTATCTGGAGGCCCTCACCGTCTGGCACCCTTAGTCA 1140
Db 1081 CAGGCCCTGGATCAGCCCTCATTTGTATCTGGAGGCCCTCACCGTCTGGCACCCTTAGTCA 1140
QY 1141 TTGAGTCAATCAGTGAATTTGCTGTTTTCTCGTCACTTTGTTCAATGGAAATTTGTTCA 1200
Db 1141 TTGAGTCAATCAGTGAATTTGCTGTTTTCTCGTCACTTTGTTCAATGGAAATTTGTTCA 1200
QY 1201 TAATATTAAAGAGAGGAGGAGGTTCAAGAGGAGCCATGGGGCACTACGCTTAGCTGAAC 1260
Db 1201 TAATATTAAAGAGAGGAGGAGGTTCAAGAGGAGCCATGGGGCACTACGCTTAGCTGAAC 1260
QY 1261 GTGAGTGACACGACGCTGCAGACTCACTGTGGGAGGAGACAAAACCTAGAGACTCAAAG 1320
Db 1261 GTGAGTGACACGACGCTGCAGACTCACTGTGGGAGGAGACAAAACCTAGAGACTCAAAG 1320
QY 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTTCAGAGAGAGTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTTCAGAGAGAGTGAACCTAAACATAGAAAT 1380

QY 1381 TGCGTCGACGAACCTCCTTGATTATTTAGCCCTTCTGTTCAATTTCTCAAAAAGATTTCGCCA 1440
Db 1381 TGCGTCGACGAACCTCCTTGATTATTTAGCCCTTCTGTTCAATTTCTCAAAAAGATTTCGCCA 1440
RESULT 3
LOCUS AR117794 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 10 from patent US 6140305.
ACCESSION AR117794
VERSION AR117794.1 GI:14098700
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnrirke,A., Ruddy,D.,
Tsuchihashi,Z. and Wolff,R.K
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 10 31-OCT-2000;
FEATURES
Location/Qualifiers
source 1. 1440
BASE COUNT 348 a 355 c 406 g 331 t
ORIGIN
Query Match 99.9%; Score 1438.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGGACACTGGATCATCCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGAGAGAGAGA 60
Db 1 GGGGACACTGGATCATCCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGAGAGAGAGA 60
QY 61 ACTAAGTTCATAAGACCTGTTGCTTTTACACAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAGTTCATAAGACCTGTTGCTTTTACACAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CTTAGGCAATAGCTGTAGGGTACCTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
Db 121 CTTAGGCAATAGCTGTAGGGTACCTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
QY 181 AAGCGGAGATTTAAGCGGACCTGCGCCAGAGCTGGGAATGGCCCGCAGCCAGGC 240
Db 181 AAGCGGAGATTTAAGCGGACCTGCGCCAGAGCTGGGAATGGCCCGCAGCCAGGC 240
QY 241 CGCGCTTCTCCTCTGATGCTTTTTCAGACCGCGCTCTGCGAGGGCGCTTGTGCGTT 300
Db 241 CGCGCTTCTCCTCTGATGCTTTTTCAGACCGCGCTCTGCGAGGGCGCTTGTGCGTT 300
QY 301 CACACTCTCTGACCTACCTCTTCATGGGTGCTCAGACGAGACCTTGGTCTTCTCTTGT 360
Db 301 CACACTCTCTGACCTACCTCTTCATGGGTGCTCAGACGAGACCTTGGTCTTCTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTTCTGCTTCTATCATGATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTTCTGCTTCTATGATCATGAGAGTCGCC 420
QY 421 GTGTGGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGCTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGCTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 AAAATCACACACACGAGGAGTCCACACCTGCAGCTCATCTGGGCTGTGAATGC 600
Db 541 AAAATCACACACACGAGGAGTCCACACCTGCAGCTCATCTGGGCTGTGAATGC 600
QY 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCAGGACCACTTG 660

Db 601 AAGAAGACACACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGCGAGGACCACTTGG 660
QY 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGGCTGGCCCAACAGC 720
Db 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAGAACCCAGGGGCTGGCCCAACAGC 720
QY 721 TGGAGTGGGAAAGGACACAGATTTCGGGCCAGGCAGAGAACAGGGCCCTACCTGGAGAGGGACT 780
Db 721 TGGAGTGGGAAAGGACACAGATTTCGGGCCAGGCAGAGAACAGGGCCCTACCTGGAGAGGGACT 780
QY 781 GCCTCTGCACAGCTGCGAGCAGTGTCTGGAGCTGGGAGAGGCTGTTTGGACCAACAAAGTCC 840
Db 781 GCCTCTGCACAGCTGCGAGCAGTGTCTGGAGCTGGGAGAGGCTGTTTGGACCAACAAAGTCC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGCGG 900
QY 901 CCTTGAACCTACTACCCCCAGACATCACCATAAGTGGCTGAAGGATTAAGCAAGCCAATGG 960
Db 901 CCTTGAACCTACTACCCCCAGACATCACCATAAGTGGCTGAAGGATTAAGCAAGCCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTAAGAGCTATTGCCAAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTAAGAGCTATTGCCAAATGGGATGGGACCTACCAAGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGAAAGCAGAGATATACGTACCAAGTGGAGCAC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGAAAGCAGAGATATACGTACCAAGTGGAGCAC 1080
QY 1081 CAGGCTGTGATCAGCCCTCATGTGATCTGGAGCCCTCACCCTGTGCGACCCCTAGTCA 1140
Db 1081 CAGGCTGTGATCAGCCCTCATGTGATCTGGAGCCCTCACCCTGTGCGACCCCTAGTCA 1140
QY 1141 TTGGAGTCAATCAGTGAATGCTGTTTTTGTGCTCATCTTGTTCATTTGGAATTTTGTCA 1200
Db 1141 TTGGAGTCAATCAGTGAATGCTGTTTTTGTGCTCATCTTGTTCATTTGGAATTTTGTCA 1200
QY 1201 TAATATTAAAGGAGGCGAGGGTTCAAGAGGAGCCATGGGCACTACGTCTTAGCTGAAC 1260
Db 1201 TAATATTAAAGGAGGCGAGGGTTCAAGAGGAGCCATGGGCACTACGTCTTAGCTGAAC 1260
QY 1261 GTGAGTGACACGAGCTGCACACTCAGTGTGGAGGAGCAGAACTAGAGACTCAAG 1320
Db 1261 GTGAGTGACACGAGCTGCACACTCAGTGTGGAGGAGCAGAACTAGAGACTCAAG 1320
QY 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCCTGACCACTCTTGAATTTAGCCTTCTCTGTTTCATTTCCCTCAAAAGATTTCCCCA 1440
Db 1381 TGCCTGACCACTCTTGAATTTAGCCTTCTCTGTTTCATTTCCCTCAAAAGATTTCCCCA 1440

RESULT 4

AR117795 LOCUS 1440 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 11 from patent US 6140305.
ACCESSION AR117795
VERSION AR117795.1 GI:14098701
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas W.J., Drayna, D.T., Feder, J.N., Gnirke, A., Ruddy, D., Tsuchihashi, Z. and Wolff, R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 11 31-Oct-2000;
FEATURES Location/Qualifiers
1..1440
/organism="unknown"

BASE COUNT 347 a 354 c 408 g 331 t
ORIGIN
Query Match 99.9%; Score 1438.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGGACACTGATCACCTAGTGTTCACAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db 1 GGGGACACTGATCACCTAGTGTTCACAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
QY 61 ACTAAAGTTCGAAAGACCTGTTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCGAAAGACCTGTTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGGCCCCCAAG 180
Db 121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGGCCCCCAAG 180
QY 181 AAGCGGAGATTTAACGGGGACCTCGGCCAGAGCTGGGAAATGGGCCCGCGAGCCAGGC 240
Db 181 AAGCGGAGATTTAACGGGGACCTCGGCCAGAGCTGGGAAATGGGCCCGCGAGCCAGGC 240
QY 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGCGCTTGTGCGTT 300
Db 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCCTGCAGGGCGCTTGTGCGTT 300
QY 301 CACACTCTCTGCACACTACTCTTCATGGGTGCTTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db 301 CACACTCTCTGCACACTACTCTTCATGGGTGCTTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTGCTGTTTCTATGATGATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCAGCTGTTGCTGTTTCTATGATGATGAGAGTCGCC 420
QY 421 GTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTCAAAAGGTGGGATCACATGTTTCACTGTTGACTTCTTGACATATTATGG 540
Db 481 TGAGTCAGAGTCTCAAAAGGTGGGATCACATGTTTCACTGTTGACTTCTTGACATATTATGG 540
QY 541 AAAATCAAAACACAGCAAGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
Db 541 AAAATCAAAACACAGCAAGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAAATGC 600
QY 601 AAGAAGACACAGTACCGAGGGCTACTGGAGTACGGGTATGATGGCAGGACCACTTG 660
Db 601 AAGAAGACACAGTACCGAGGGCTACTGGAGTACGGGTATGATGGCAGGACCACTTG 660
QY 661 AATTCTGCCCTGCACACTGGATTGGAGAGCAGCAGAACCCAGGGCTGGCCCAACAGC 720
Db 661 AATTCTGCCCTGCACACTGGATTGGAGAGCAGCAGAACCCAGGGCTGGCCCAACAGC 720
QY 721 TGGAGTGGGAAAGGACACAAGATTTCGGGCCAGGCAGAGAACAGGGCCCTACCTGGAGAGG 780
Db 721 TGGAGTGGGAAAGGACACAAGATTTCGGGCCAGGCAGAGAACAGGGCCCTACCTGGAGAGG 780
QY 781 GCGCTGCACAGCTGCACAGCTTGTGAGGCTGGGAGAGGTTGTTGGACCAACAAGTGC 840
Db 781 GCGCTGCACAGCTGCACAGCTTGTGAGGCTGGGAGAGGTTGTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
QY 901 CCTTGAACCTACTACCCCCAGACATCACCATAAGTGGCTGAAGGATTAAGCAAGCCAATGG 960
Db 901 CCTTGAACCTACTACCCCCAGACATCACCATAAGTGGCTGAAGGATTAAGCAAGCCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTAAGAGCTATTGCCAAATGGGATGGGACCTACCAAGGCT 1020

Db 961 ATGCCAAGGAGTTGCAACCTAAAGACGTATTGCCAATGGGATGGGACCTACCGAGGCT 1020
QY 1021 GGATAACCTTTGGGCTGTACCCCTGGGGAAGACAGAGATATACGTACCAAGTGGAGCACC 1080
Db 1021 GGATAACCTTTGGGCTGTACCCCTGGGGAAGACAGAGATATACGTCCAGGTGGAGCACC 1080
QY 1081 CAGGCTGGATACAGCCCTCATTTGTGATCTGGAGCCCTACCGCTGTGGACCCCTAGTCA 1140
Db 1081 CAGGCTGGATACAGCCCTCATTTGTGATCTGGAGCCCTACCGCTGTGGACCCCTAGTCA 1140
QY 1141 TTGAGTCATCAGTGGAAATTCGTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA 1200
Db 1141 TTGAGTCATCAGTGGAAATTCGTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA 1200
QY 1201 TAATATTAAAGAAAGAGCGAGGGTTCAAGAGAGCCATGGGCACTACGCTTTAGCTGAAC 1260
Db 1201 TAATATTAAAGAAAGAGCGAGGGTTCAAGAGAGCCATGGGCACTACGCTTTAGCTGAAC 1260
QY 1261 GTGAGTCACACGAGCTGCAGACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTCACACGAGCTGCAGACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320
QY 1321 AGGAGTGCATTTATGAGCTTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCCTGACGACACTCCTTTAGCTTCTCTGTTTCATTTCTCTGTTTCATTTCTCAAAAAGATTTCCCA 1440
Db 1381 TGCCTGACGAACTCCTTTGATTTAGCTTCTCTGTTTCATTTCTCTGTTTCATTTCTCAAAAAGATTTCCCA 1440

RESULT 5
AR149464
LOCUS AR149464 1440 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 10 from patent US 6228594.
ACCESSION AR149464
VERSION AR149464.1 GI:15114055
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuehishashi,Z. and Wolff,R.K.
TITLE Method for determining the presence or absence of a hereditary
hemochromatosis gene mutation
JOURNAL Patent: US 6228594-A 10 08-MAY-2001;
FEATURES
source Location/Qualifiers
1..1440
BASE COUNT 348 a 355 c 406 g 331 t
ORIGIN

Query Match 99.9%; Score 1438.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
Db 1 GGGGACACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGAGA 60
QY 61 ACTAAAGTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTCTGAAAGACCTGTTGCTTTTCACCAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CCTAGGCAATAGCTGTAGGTGACTCTGAGGCCATCCCGGTTTCCCGCCCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGTAGGTGACTCTGAGGCCATCCCGGTTTCCCGCCCCCAAAAG 180
QY 181 AAGCGGAGATTTACGGGACGTGGCCAGAGCTGGGAAATGGCCCGGAGCCAGGC 240

Db 181 AAGCGGAGATTTAAGCGGACGTGGGCCACAGAGCTGGGGAATGGCCCGGAGCCAGGC 240
QY 241 CGGCGCTTCTCTCTCTGATGCTTTTTCAGACCGCGGTCCTTCAGGGGCGCTTGTGCGTT 300
Db 241 CGGCGCTTCTCTCTCTGATGCTTTTTCAGACCGCGGTCCTTCAGGGGCGCTTGTGCGTT 300
QY 301 CACACTCTCTGACACTACCTCTTATGGGTGCTTCAGACAGACCTTGGTCTTTCCTTGT 360
Db 301 CACACTCTCTGACACTACCTCTTATGGGTGCTTCAGACAGACCTTGGTCTTTCCTTGT 360
QY 361 TTGAAGCTTTTGGGCTACGTGGATGACAGCTGTTCTGCTTCTATGATCATGAGAGTCGCC 420
Db 361 TTGAAGCTTTTGGGCTACGTGGATGACAGCTGTTCTGCTTCTATGATCATGAGAGTCGCC 420
QY 421 GTGGAGCCCGGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGGAGCCCGGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTTCTGCTTCTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTTCTGCTTCTGACTTCTGGACTATTATGG 540
QY 541 AAAATCAACAACACACAGAGAGTCCCACACCCCTCAGGTCATCTCGGCTGTGAAATGC 600
Db 541 AAAATCAACAACACACAGAGAGTCCCACACCCCTCAGGTCATCTCGGCTGTGAAATGC 600
QY 601 AAGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGACACCACTTG 660
Db 601 AAGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGACACCACTTG 660
QY 661 AATTCTGCCCTCACACACTGGATTGGAGAGCAGACAGACCCAGGSCCTGGCCACCAAGC 720
Db 661 AATTCTGCCCTCACACACTGGATTGGAGAGCAGACAGACCCAGGSCCTGGCCACCAAGC 720
QY 721 TGGAGTGGGAAAGGCACAAAGATTTCGGGCCAGCAGAACAGGSCCTTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAAGGCACAAAGATTTCGGGCCAGCAGAACAGGSCCTTACCTGGAGAGGACT 780
QY 781 GCCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGAGTGTTTTGACCAACAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGAGTGTTTTGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGCACATCATGTACCTCTTCAGTGACCACTCTACGGTGTCCGG 900
Db 841 CTCCTTTGGTGAAGGTGCACATCATGTACCTCTTCAGTGACCACTCTACGGTGTCCGG 900
QY 901 CCTTGAACCTACTACCCCAAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
Db 901 CCTTGAACCTACTACCCCAAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
QY 961 ATGCCAAGGAGTTGAACTTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGGAGTTGAACTTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020
QY 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACCAAGTGGAGCACC 1080
Db 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACCAAGTGGAGCACC 1080
QY 1081 CAGGCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCCTGTGGACCCCTAGTCA 1140
Db 1081 CAGGCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCCTGTGGACCCCTAGTCA 1140
QY 1141 TTGAGTCATCAGTGGAAATTCGTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA 1200
Db 1141 TTGAGTCATCAGTGGAAATTCGTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA 1200
QY 1201 TAATATTAAAGAAAGAGCGAGGGTTCAAGAGAGCCATGGGCACTACGCTTTAGCTGAAC 1260
Db 1201 TAATATTAAAGAAAGAGCGAGGGTTCAAGAGAGCCATGGGCACTACGCTTTAGCTGAAC 1260
QY 1261 GTGAGTCACACGAGCTGCAGACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTCACACGAGCTGCAGACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320

QY 1321 AGGAGTGCATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
|||||
Db 1321 AGGAGTGCATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
|||||
QY 1381 TGCTGACGACACTCCTTGATTTTAGCCTCTCTGTTTCATTTCTCTCAAAAAGATTTCCCCA 1440
|||||
Db 1381 TGCTGACGACACTCCTTGATTTTAGCCTCTCTGTTTCATTTCTCTCAAAAAGATTTCCCCA 1440
|||||
RESULT 6
AR149465 AR149465 1440 bp DNA linear PAT 08-AUG-2001
LOCUS Sequence 11 from patent US 6228594.
DEFINITION AR149465
ACCESSION AR149465
VERSION AR149465.1 GI:15114056
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuchihashi,Z. and Wolff,R.K.
TITLE Method for determining the presence or absence of a hereditary
hemochromatosis gene mutation
JOURNAL Patent: US 6228594-A 11 08-MAY-2001;
FEATURES Location/Qualifiers
source 1..1440
BASE COUNT 347 a 354 c 408 g 331 t
ORIGIN
Query Match 99.9%; Score 1438.4; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGGGACACTGGATACACTAGTGTGTTTCAAGCAGGTAAGTCTGCTGTAGGAGAGAGA 60
Db 1 GGGGACACTGGATACACTAGTGTGTTTCAAGCAGGTAAGTCTGCTGTAGGAGAGAGA 60
QY 61 ACTAAAGTTCTGAAGACCTGTTGCTTTTCCACAGAGAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCTGAAGACCTGTTGCTTTTCCACAGAGAGTTTACTGGGCATCTCCTGAG 120
QY 121 CTAGGCAATAGCTGTAGGTTGACTTCTGGAGCCATCCCGCTTTCCCGCCCCCAAAAG 180
Db 121 CTAGGCAATAGCTGTAGGTTGACTTCTGGAGCCATCCCGCTTTCCCGCCCCCAAAAG 180
QY 181 AAGCGAGATTTAAGCGGACCTGGGGCAGAGCTGGGGAATGGCGCGGAGCCAGGC 240
Db 181 AAGCGGAGATTTAAGCGGACCTGGGGCAGAGCTGGGGAATGGCGCGGAGCCAGGC 240
QY 241 CGCGCGTCTCTCTCTGATGCTTTTGCAGACCGCGGTCTGTCAGGGCGCTTGTGCGCTT 300
Db 241 CGCGCGTCTCTCTCTGATGCTTTTGCAGACCGCGGTCTGTCAGGGCGCTTGTGCGCTT 300
QY 301 CACACTCTCTGACACTACCTCTTCATGGGTGCTCAGACAGGACCTTGSTCTTCCTTGT 360
Db 301 CACACTCTCTGCACTACCTCTTCATGGGTGCTCAGACAGGACCTTGSTCTTCCTTGT 360
QY 361 TTGACGTTTGGGCTACGTTGGATGACCAGCTGTTGCTGTTCTATGATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTTGGATGACCAGCTGTTGCTGTTCTATGATGAGAGTCGCC 420
QY 421 GTGTGAGCGCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGAGCGCCGAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540

QY 541 AAAATCAACAACCACAGCAAGGAGTCCACACCCCTGCAGGTTCATCTGGGTGTGAAATGC 600
|||||
Db 541 AAAATCAACAACCACAGCAAGGAGTCCACACCCCTGCAGGTTCATCTGGGTGTGAAATGC 600
|||||
QY 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGCAGACCACTTG 660
|||||
Db 601 AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGCAGACCACTTG 660
|||||
QY 661 AATTCTGCCCTGCACACTGATTGGAGAGCAGACCAAGCCAGGGCTTGCCCCCACCAGC 720
|||||
Db 661 AATTCTGCCCTGCACACTGATTGGAGAGCAGACCAAGCCAGGGCTTGCCCCCACCAGC 720
|||||
QY 721 TGGAGTGGAAAGGCACAAGATTTCGGGCCAGGACAGACAGGGCTTACTCTGGAGAGGACT 780
|||||
Db 721 TGGAGTGGAAAGGCACAAGATTTCGGGCCAGGACAGACAGGGCTTACTCTGGAGAGGACT 780
|||||
QY 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
|||||
Db 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840
|||||
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGCACACACTACGGTGTCTGGG 900
|||||
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGCACACACTACGGTGTCTGGG 900
|||||
QY 901 CTTTGAAGTACTACCCCGCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
|||||
Db 901 CTTTGAAGTACTACCCCGCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
|||||
QY 961 ATGCCAAGGAGTTTCGAACCTTAAGACGTTATGCCCATGGGATGGGACCTACGAGGGCT 1020
|||||
Db 961 ATGCCAAGGAGTTTCGAACCTTAAGACGTTATGCCCATGGGATGGGACCTACGAGGGCT 1020
|||||
QY 1021 GGATAACCTTGGCTGTACCCCTCGGGAAGCAGAGATATACGTACACAGGTGGAGCACC 1080
|||||
Db 1021 GGATAACCTTGGCTGTACCCCTCGGGAAGCAGAGATATACGTACACAGGTGGAGCACC 1080
|||||
QY 1081 CAGCCCTGGATCAGCCCTCATTTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA 1140
|||||
Db 1081 CAGCCCTGGATCAGCCCTCATTTGATCTGGGAGCCCTCACCGTCTGGCACCCCTAGTCA 1140
|||||
QY 1141 TTGAGTCACTCAGTGAATGCTGTTTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA 1200
|||||
Db 1141 TTGAGTCACTCAGTGAATGCTGTTTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA 1200
|||||
QY 1201 TAATATTAAAGGAAGCAGGCTTCAAGAGAGCCATGGGGCCTACGTCTTACGTGAAC 1260
|||||
Db 1201 TAATATTAAAGGAAGCAGGCTTCAAGAGAGCCATGGGGCCTACGTCTTACGTGAAC 1260
|||||
QY 1261 GTGAGTGACAGCGAGCTGCAGACTCACTGTGGGAAGGACAAAACCTAGAGACTCAAG 1320
|||||
Db 1261 GTGAGTGACAGCGAGCTGCAGACTCACTGTGGGAAGGACAAAACCTAGAGACTCAAG 1320
|||||
QY 1321 AGGAGTGCATTTATGAGCTTTCATGTTTCAGGAGAGTTTGAACCTAAACATAGAAAT 1380
|||||
Db 1321 AGGAGTGCATTTATGAGCTTTCATGTTTCAGGAGAGTTTGAACCTAAACATAGAAAT 1380
|||||
QY 1381 TGCTGACGAACTCCTTGATTTAGCCCTTCTGTTTCATTTCTCAAAAAGATTTCCCCA 1440
|||||
Db 1381 TGCTGACGAACTCCTTGATTTAGCCCTTCTGTTTCATTTCTCAAAAAGATTTCCCCA 1440
|||||

RESULT 7

AR117793 AR117793 1440 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 9 from patent US 6140305.
DEFINITION AR117793
ACCESSION AR117793
VERSION AR117793.1 GI:14098699
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,

Tsuchihashi, Z. and Wolff, R.K.
Hereditary hemochromatosis gene products
Patent: US 6140305-A 9 31-OCT-2000;
Location/Qualifiers
1..1440
/organism="unknown"
BASE COUNT 347 a 355 c 407 g 331 t
ORIGIN

Query Match 99.8%; Score 1436.8; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTGAGGAGAGAGA 60
DB 1 GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTGAGGAGAGAGA 60
QY 61 ACTAAGTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
DB 61 ACTAAGTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CTTAGGCAATAGCTGTAGGGTACCTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
DB 121 CTTAGGCAATAGCTGTAGGGTACCTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
QY 181 AAGCGGAGATTTAAGCGGACGTGGGCCAGAGCTGGGGAATGGGCCGAGCCAGGC 240
DB 181 AAGCGGAGATTTAAGCGGACGTGGGCCAGAGCTGGGGAATGGGCCGAGCCAGGC 240
QY 241 CGGCGTCTCTCTCTGATGCTTTTGCAGACCGCGGTCTCGAGGGGGCTTGTGCGGTT 300
DB 241 CGGCGTCTCTCTCTGATGCTTTTGCAGACCGCGGTCTCGAGGGGGCTTGTGCGGTT 300
QY 301 CACACTCTCTGACTACCTCTTATGCGGTGCTCAGAGCAGGACCTTGGTCTTCTCTGT 360
DB 301 CACACTCTCTGACTACCTCTTATGCGGTGCTCAGAGCAGGACCTTGGTCTTCTCTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTCTGTTCTATGATGATGAGTCGCC 420
DB 361 TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTCTGTTCTATGATGATGAGTCGCC 420
QY 421 GTGTGGAGCCCCGAACCTCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
DB 421 GTGTGGAGCCCCGAACCTCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
DB 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 AAAATCACACACACAGCAGGAGTCCACACCCCTGCAGGTATCTCTGGGCTGTGAAATGC 600
DB 541 AAAATCACACACACAGCAGGAGTCCACACCCCTGCAGGTATCTCTGGGCTGTGAAATGC 600
QY 601 AAGAAGACACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
DB 601 AAGAAGACACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
QY 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGACCCAGGGCTGGGCCACCAAGC 720
DB 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGACCCAGGGCTGGGCCACCAAGC 720
QY 721 TGGAGTGGGAAGGCAACAAGATTGGGCCAGGCAGAACAGGCGCTACCTGGAGAGGGACT 780
DB 721 TGGAGTGGGAAGGCAACAAGATTGGGCCAGGCAGAACAGGCGCTACCTGGAGAGGGACT 780
QY 781 GCCTTCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTTTGGACCAACAGTGC 840
DB 781 GCCTTCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTTTGGACCAACAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTCTACGGGTGCGGG 900
DB 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTCTACGGGTGCGGG 900

RESULT 8

AR149463
LOCUS AR149463 Sequence 9 from patent US 6228594. 1440 bp DNA linear PAT 08-AUG-2001
DEFINITION AR149463
ACCESSION AR149463
VERSION AR149463.1 GI:15114054
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1440)
AUTHORS Thomas, W.J., Drayna, D.T., Feder, J.N., Gnirke, A., Ruddy, D., Tsuchihashi, Z. and Wolff, R.K.
TITLE Method for determining the presence or absence of a hereditary hemochromatosis gene mutation
JOURNAL Patent: US 6228594-A 9 08-MAY-2001;
FEATURES Location/Qualifiers
source 1..1440
BASE COUNT 347 a 355 c 407 g 331 t
ORIGIN

Query Match 99.8%; Score 1436.8; DB 6; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTGAGGAGAGAGA 60
DB 1 GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTGAGGAGAGAGA 60
QY 61 ACTAAGTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
DB 61 ACTAAGTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120

QY 121 CCTAGGCAATAGCTAGGTGAGTCTCTGGAGCCATCCCGCTTTCGCCGCCCCCAAAAG 180
Db 121 CCTAGGCAATAGCTAGGTGAGTCTCTGGAGCCATCCCGCTTTCGCCGCCCCCAAAAG 180
QY 181 AAGCGGAGATTTAAACGGGACGTCGGGCCAGAGCTGGGGAATGGCGCCGCGAGCGC 240
Db 181 AAGCGGAGATTTAAACGGGACGTCGGGCCAGAGCTGGGGAATGGCGCCGCGAGCGC 240
QY 241 CGGCGCTTTCCTCCCTGATGCTTTTCAGACCGCGCTCTCGAGGGCGCTTCTGCGCTT 300
Db 241 CGGCGCTTTCCTCCCTGATGCTTTTCAGACCGCGCTCTCGAGGGCGCTTCTGCGCTT 300
QY 301 CACACTCTGCACTACCTCTTCATGGGTGCGCTCAGAGCAGGACCTTGTCTTCTTGT 360
Db 301 CACACTCTGCACTACCTCTTCATGGGTGCGCTCAGAGCAGGACCTTGTCTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACACGCTGTTCTGCTTCTATGATGATGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACACGCTGTTCTGCTTCTATGATGATGAGTCGCC 420
QY 421 GTGTGAGCCCCGAATCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGAGCCCCGAATCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGACTCTGAAGGGTGGATCACATGTTCACTGTTTCACTTCTGCACTATTATGG 540
Db 481 TGAGTCAGACTCTGAAGGGTGGATCACATGTTCACTGTTTCACTTCTGCACTATTATGG 540
QY 541 AAATCACACACACAGCAAGGAGTCCACACCCCTGCAGGTCACTCTGGGCTGTGAATGC 600
Db 541 AAATCACACACACAGCAAGGAGTCCACACCCCTGCAGGTCACTCTGGGCTGTGAATGC 600
QY 601 AAGAACAACAGTACCGAGGCTACTGGAATACGGGTATGATGGGAGGACCACTTG 660
Db 601 AAGAACAACAGTACCGAGGCTACTGGAATACGGGTATGATGGGAGGACCACTTG 660
QY 661 AATTCTGCGCTGACACACTTGGATGGAGAGCAGACCCAGGCGCTGGCCCAACAGC 720
Db 661 AATTCTGCGCTGACACACTTGGATGGAGAGCAGACCCAGGCGCTGGCCCAACAGC 720
QY 721 TGGAGTGGGAAGGCAACAGATTCGGGCCAGCAGAACAGGCGCTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAGGCAACAGATTCGGGCCAGCAGAACAGGCGCTACCTGGAGAGGACT 780
QY 781 GCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGGCTTTTGGACCAACAGTGC 840
Db 781 GCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGAGAGGCTTTTGGACCAACAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
QY 901 CTTGAACTACTACCCCAAGACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
Db 901 CTTGAACTACTACCCCAAGACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTTAAGACGATTTGCCCAATGGGGATGGGACCTACAGGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTTAAGACGATTTGCCCAATGGGGATGGGACCTACAGGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACCAGTGGAGCAC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACCAGTGGAGCAC 1080
QY 1081 CAGGCGTGTATCAGCCCTCATTTGATCTGGGAGCCCTCACCGCTGGCACCCCTAGTCA 1140
Db 1081 CAGGCGTGTATCAGCCCTCATTTGATCTGGGAGCCCTCACCGCTGGCACCCCTAGTCA 1140
QY 1141 TTGGAGTATCAGTGAATGCTGTTTTCGTCATCTTGTTCATTTGGAATTTTGTTC 1200
Db 1141 TTGGAGTATCAGTGAATGCTGTTTTCGTCATCTTGTTCATTTGGAATTTTGTTC 1200
QY 1201 TAATATTAAGGAAGAGCGGGGTTCAAGAGGAGCCATGGGCACTACCTCTTAGCTGAAC 1260

Db 1201 TAATATTAAGGAAGAGCGGGTTCAGAGAGCCATGGGCACTACCTCTTAGCTGAAC 1260
QY 1261 GTGAGTCACAGCGAGCCTGACAGACTCTGTGGGAAGGAGACAAAAGTAGAGACTCAAAG 1320
Db 1261 GTGAGTCACAGCGAGCCTGACAGACTCTGTGTGGGAAGGAGACAAAAGTAGAGACTCAAAG 1320
QY 1321 AGGAGTGCATTTATGAGCTCTTCATGCTTTTCAGGAGAGCTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTCTTCATGCTTTTCAGGAGAGCTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCTGACGAACCTCCTTGATTTAGCCTTCTCTCTTCATTTCCCAAAAAGATTTCCTCCA 1440
Db 1381 TGCTGACGAACCTCCTTGATTTAGCCTTCTCTCTTCATTTCCCAAAAAGATTTCCTCCA 1440
RESULT 9
HSU60319 HSU60319 2727 bp mRNA linear PRI 29-OCT-1997
LOCUS Homo sapiens haemochromatosis protein (HLA-H) mRNA, complete cds.
DEFINITION U60319
ACCESSION U60319
VERSION U60319.1 GI:1469789
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2727)
AUTHORS Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A.,
Basava,A., Dornishian,F., Domingo,R., Ellis,M.C., Fullan,A.,
Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P.,
Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C.,
Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E.,
Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J.,
Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.,
A novel MHC class I-like gene is mutated in patients with
hereditary haemochromatosis
Nature Genet. 13 (4), 399-408 (1996)
JOURNAL 96331279
MEDLINE
REFERENCE 2 (bases 1 to 2727)
AUTHORS Feder,J.N., Gnirke,A., Thomas,W., Tsuchihashi,Z., Ruddy,D.A.,
Basava,A., Dornishian,F., Domingo,R., Ellis,M.C., Fullan,A.,
Hinton,L.M., Jones,N.L., Kimmel,B.E., Kronmal,G.S., Lauer,P.,
Lee,V.K., Loeb,D.B., Mapa,F., McClelland,E., Meyer,N.C.,
Mintier,G.A., Moeller,N., Moore,T., Morkang,E., Prass,C.E.,
Quintana,L., Stranes,S.M., Schatzman,R.C., Brunke,K.J.,
Drayna,D.T., Risch,N.J., Bacon,B.R. and Wolff,R.K.,
Direct Submission
TITLE
JOURNAL Submitted (10-JUN-1996) Mercator Genetics, 4040 Campbell Ave.,
Menlo Park, CA 94025, USA
FEATURES
source
1. .2727
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
1. .2727
/gene="HLA-H"
/note="HFE"
222. .1268
/gene="HLA-H"
/codon_start=1
/product="haemochromatosis protein"
/protein_id="AAC51823.1"
/db_xref="GI:1469790"
/translation="MGPRAPALLLLMLLOTAVLOGRLLRSHSLHLPFGASEODLGL
SLFEALGYVDDQLFVFDHESRVERPTFWSSRISQNWQLQSLSLGWDHMTVD
WTIMENHNKSHESHTLQVLGCEMDNSTEGYWKYDGDQDHLFECDPTLDWRAAEP
RAWPKLEWERHKIRARONRAYLERDQPAQLLEGLVDQVPLVPLVKTHTVTS
SVTLTRCALNYTPQNTMKWLKQPMDFKPEKPLNGDGTGYQGWITLAVPPGE
EORYTCQVEHPGLDQPLIWIWEPSPSGTLIVIGVISIAVFFVILFILTILRKRG
SRGANGHYVLAERE"
BASE COUNT 702 a 606 c 660 g 759 t

ORIGIN

Query Match 99.8%; Score 1436.8; DB 9; Length 2727;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0;

Qy	1	GGGACACTGGATCACTAGTGTTCACAAAGCAGGTACCTTCTGCTGTAGAGAGAGAGA	60
Db	1	GGGACACTGGATCACTAGTGTTCACAAAGCAGGTACCTTCTGCTGTAGAGAGAGAGA	60
Qy	61	ACTAAAGTCTCGAAAGACCTGTGTCTTTTCACCAAGGAAGTTTTACTGGGCATCTCCTGAG	120
Db	61	ACTAAAGTCTCGAAAGACCTGTGTCTTTTCACCAAGGAAGTTTTACTGGGCATCTCCTGAG	120
Qy	121	CCTAGGCAATPAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCGCCGCCCAAAAG	180
Db	121	CCTAGGCAATPAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCGCCGCCCAAAAG	180
Qy	181	AAGCGGAGATTTAACGGGACGTGCGGCAGAGCTGGGGAAATGGCCCGCGAGCCAGGC	240
Db	181	AAGCGGAGATTTAACGGGACGTGCGGCAGAGCTGGGGAAATGGCCCGCGAGCCAGGC	240
Qy	241	CGGCGCTTCTCCCTCCTGATGCTTTTGCAGACCGCGGTCTCTGAGGGGCGCTTGCTGGGTT	300
Db	241	CGGCGCTTCTCCCTCCTGATGCTTTTGCAGACCGCGGTCTCTGAGGGGCGCTTGCTGGGTT	300
Qy	301	CACACTCTCTGCACTACCTCTCTCATNGGGTGCCCTCAGACGAGACCTTGGTCTTTCCTTGT	360
Db	301	CACACTCTCTGCACTACCTCTCTCATNGGGTGCCCTCAGACGAGACCTTGGTCTTTCCTTGT	360
Qy	361	TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTCGTGTCTTATGATCATGAGAGTGC	420
Db	361	TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTCGTGTCTTATGATCATGAGAGTGC	420
Qy	421	GTGPGAGCCCCGAATCCCATGCGTTTCCAGTAGAATTTCAAGCCAGATGTGGGTGCAGC	480
Db	421	GTGPGAGCCCCGAATCCCATGCGTTTCCAGTAGAATTTCAAGCCAGATGTGGGTGCAGC	480
Qy	481	TGAGTCAGAGCTGTGAAGGGTGGGATCACATGTTTCACCTGTTGACPTCTGGACTATTATGG	540
Db	481	TGAGTCAGAGCTGTGAAGGGTGGGATCACATGTTTCACCTGTTGACPTCTGGACTATTATGG	540
Qy	541	AAATTCACACACCACAGCAAGGAGTCCACACCCCTGCAGGTCTATCTTGGGCTGTGAAATGC	600
Db	541	AAATTCACACACCACAGCAAGGAGTCCACACCCCTGCAGGTCTATCTTGGGCTGTGAAATGC	600
Qy	601	AAGAAGCAACACTACCGAGGCTACTGGAAGTACGGGTATGATGGCGAGGACCACTTG	660
Db	601	AAGAAGCAACACTACCGAGGCTACTGGAAGTACGGGTATGATGGCGAGGACCACTTG	660
Qy	661	AATTCTGCCCTGCACACTGTGATTGGAGAGCAGCAGAACCCAGGGCCTTGGCCCAACCAAGC	720
Db	661	AATTCTGCCCTGCACACTGTGATTGGAGAGCAGCAGAACCCAGGGCCTTGGCCCAACCAAGC	720
Qy	721	TGGAGTGGGAAAGCACAAGATTTCGGGCCAGGCAAGACAGGGCCTACCTGGAGAGGACT	780
Db	721	TGGAGTGGGAAAGCACAAGATTTCGGGCCAGGCAAGACAGGGCCTACCTGGAGAGGACT	780
Qy	781	GCCCTGCACAGCTGCAGCAGTCTCTGGAGCTGGGAGAGGTCTTTTGGACCAACAAAGTGC	840
Db	781	GCCCTGCACAGCTGCAGCAGTCTCTGGAGCTGGGAGAGGTCTTTTGGACCAACAAAGTGC	840
Qy	841	CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGGTGCGGG	900
Db	841	CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGGTGCGGG	900
Qy	901	CCTTGAACACTACTACCCCCGAACATCACCATGAAGTGGCTGAAGATGAAGCAGCAATGG	960
Db	901	CCTTGAACACTACTACCCCCGAACATCACCATGAAGTGGCTGAAGATGAAGCAGCAATGG	960
Qy	961	ATGCCAAGGAGTTGCAACCTTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT	1020

[illegible]

VEPRTPWVSSRISSQMWLQSLKGMHMFVDFWTFMINHNSKESHTLIQVILGCE
MOEDNSTEYKYGVDGQHLFEFQPDITLDRWAEPRAWPTKLEWRKIRARONRAYL
ERDCPAOLQLELGRGVLDQOVPPLVKTVTHVTSVTLRLRALNYVPONTMKNWLK
DKOPMDAKEFEFPKVLNPGDGTGOGWITLAVPGEBORYTCQVEHPGLDPLIWIWEP
SPSGTLVIGVISIAVFFVILFIFILFKRGRSGAMGHYVLAERE"

BASE COUNT 311 a 314 c 371 g 284 t
ORIGIN

Query Match 83.2%; Score 1197.8; DB 9; Length 1280;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 1278; Conservative 0; Mismatches 2; Indels 69; Gaps 1;

QY	62	CTAAAGTCTCGAAGACCTGTGCTTTTCCACGAGGAAGTTTACTGGGCATCTCCTGAGC	121
DB	1	CTAAAGTCTCGAAGACCTGTGCTTTTCCACGAGGAAGTTTACTGGGCATCTCCTGAGC	60
QY	122	CTAGGCAANTAGCTGTAGGTGTACTCTTGAGGCCATCCCGTTTCCCGCCCCCAAGA	181
DB	61	CTAGGCAANTAGCTGTAGGTGTACTCTTGAGGCCATCCCGTTTCCCGCCCCCAAGA	120
QY	182	AGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATGGCCCGCAGCCAGGCC	241
DB	121	AGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATGGCCCGCAGCCAGGCC	180
QY	242	GGCGCTTCTCCTCGTATGCTTTTGCAGACCGCGTCTCTGCAAGGGCGCTTCTGCTGTC	301
DB	181	GGCGCTTCTCCTCGTATGCTTTTGCAGACCGCGTCTCTGCAAGGGCGCTTCTGCTGTC	236
QY	302	ACACTCTCTGCACACTACCTCTTCATGGGTGCCTCAGACGAGGACCTTGGTCTTCTGTT	361
DB	237	-----	236
QY	362	TGAAGCTTTGGGCTACGTGGATGACAGCTGTTCTGTTCTATGATGATGAGAGTCGCGC	421
DB	237	-----CTTTGGGCTACGTGGATGACAGCTGTTCTGTTCTATGATGATGAGAGTCGCGC	291
QY	422	TGTGGAGCCCGAATCCATNGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTCGAGCT	481
DB	292	TGTGGAGCCCGAATCCATNGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTCGAGCT	351
QY	482	GAGTCAGACTCTGAAGGGTGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGA	541
DB	352	GAGTCAGACTCTGAAGGGTGGATCACATGTTCACTGTTGACTTCTGGACTATTATGGA	411
QY	542	AAATCAACACCACAGCAAGGAGTCCACACCCCTGCAGGTCTATCTGGGCTGTGAATGCA	601
DB	412	AAATCAACACCACAGCAAGGAGTCCACACCCCTGCAGGTCTATCTGGGCTGTGAATGCA	471
QY	602	AGAAGACAACAGTACCGAGGGTACTTGAAGTACGGGTATGATGGCAGGACCACTTGA	661
DB	472	AGAAGACAACAGTACCGAGGGTACTTGAAGTACGGGTATGATGGCAGGACCACTTGA	531
QY	662	ATTCTGCCCTGACACACTGGATTGAGAGCAGCAGAACCCAGGGCTGGCCCAACCACT	721
DB	532	ATTCTGCCCTGACACACTGGATTGAGAGCAGCAGAACCCAGGGCTGGCCCAACCACT	591
QY	722	GGAGTGGGAAAGCCAAAGATTTCGGGCCAGGCAACACAGGGCTACCTTGGAGAGGAGCTG	781
DB	592	GGAGTGGGAAAGCCAAAGATTTCGGGCCAGGCAACACAGGGCTACCTTGGAGAGGAGCTG	651
QY	782	CCCTGACAGCTGACAGCTAGTCTGAGCTGGGGAGAGGTGTTTGGACCAACAAGTCCC	841
DB	652	CCCTGACAGCTGACAGCTAGTCTGAGCTGGGGAGAGGTGTTTGGACCAACAAGTCCC	711
QY	842	TCCCTTGGTGAAGGTGACACATCATGTGACCTTCTCAGTGACCCTCTACCGTGTCCGGC	901
DB	712	TCCCTTGGTGAAGGTGACACATCATGTGACCTTCTCAGTGACCCTCTACCGTGTCCGGC	771
QY	902	CTTGAACCTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGGA	961
DB	772	CTTGAACCTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCAATGGA	831

QY	952	TGCCAAGAGTTCGAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCTG	1021
DB	832	TGCCAAGAGTTCGAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCTG	891
QY	1022	GATAACCTTGGCTGTACCCCTGGGGAAGACGAGATATACGTACCAAGTGGAGCACCC	1081
DB	892	GATAACCTTGGCTGTACCCCTGGGGAAGACGAGATATACGTACCAAGTGGAGCACCC	951
QY	1082	AGGCTCGATCAGCCCTCATTTGATCTGGGACCCCTACCGCTGCGACCCCTAGTCAT	1141
DB	952	AGGCTCGATCAGCCCTCATTTGATCTGGGACCCCTACCGCTGCGACCCCTAGTCAT	1011
QY	1142	TGGAGTCATCAGTGAATTCGCTTTTGTGCTCATCTTGTTCATTTGGAATTTTGTTCAT	1201
DB	1012	TGGAGTCATCAGTGAATTCGCTTTTGTGCTCATCTTGTTCATTTGGAATTTTGTTCAT	1071
QY	1202	AATATTAGGAAGAGGAGGTTTCAAGAGGAGCCATGGGCACTACGCTTAGTGAACG	1261
DB	1072	AATATTAGGAAGAGGAGGTTTCAAGAGGAGCCATGGGCACTACGCTTAGTGAACG	1131
QY	1262	TGAGTGACACGACCCCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAGA	1321
DB	1132	TGAGTGACACGACCCCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAGA	1191
QY	1322	GGAGTCATTTATGAGCTCTTTCATGTTTCAGGAGAGAGTTGAACCTTAAACATAGAATT	1381
DB	1192	GGAGTCATTTATGAGCTCTTTCATGTTTCAGGAGAGAGTTGAACCTTAAACATAGAATT	1251
QY	1382	GCCTGACGAACTCCTTGATTTAGCCCTTC	1410
DB	1252	GCCTGACGAACTCCTTGATTTAGCCCTTC	1280

RESULT 11

AF115265	AF115265	1200 bp	mRNA	linear	PRI 07-MAY-2001
LOCUS	Homo sapiens hemochromatosis termination variant terE6 (HFE) mRNA,				
DEFINITION	complete cds.				
ACCESSION	AF115265				
VERSION	AF115265.1 GI:11094314				
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1200)				
AUTHORS	Thenie,A., Orhant,M., Gicquel,I., Fergelot,P., Le Gall,J.Y., David,V. and Mosser,J.				
TITLE	The HFE gene undergoes alternate splicing processes				
JOURNAL	Blood Cells Mol. Dis. 26 (2), 155-162 (2000)				
MEDLINE	20448010				
PUBMED	11001625				
REFERENCE	2 (bases 1 to 1200)				
AUTHORS	Thenie,A., Orhant,M. and Mosser,J.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-DEC-1998) UPR 41 CNRS, Faculte de Medecine, 2, av du Pr. Bernard, Rennes 35043, France				
FEATURES	Location/Qualifiers				
source	1..1200				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="6"				
	/map="6p22"				
gene	1..1200				
	/gene="HFE"				
CDS	12..1058				
	/codon_start=1				
	/product="hemochromatosis termination variant terE6"				
	/protein_id="AAG29572.1"				
	/db_xref="GI:11094315"				
	/translation="MGPRARPAALLMLLQTAAYLOGRLLRSHLSHYLPMWASEQDLGL				
	SLFEALGYVDDQLFVFYDHSRRVETPTWSSRSQMWLQSLKGMHMFVDF				

WTIMENHNHKSHTLQVILGCEMDENSTEGYWKYGYDGDHLEFCDPDLTLDWRAAEP
RAMPTKLEWERHRIARONRAYLERDPAQLLELGLGVLDQVPLVKTTHVTS
SVYTLRCALNYYPONITMKWLKQPMDAKEFEKDPNGDGYOGWITLAVPPGE
EORYTCOVHEPGLDPLVWEPSPSLVIGVISGIAVFVILFIFILRKRG
SRGANGHYVLAERE"

BASE COUNT	298 a	290 c	346 g	266 t	
ORIGIN					
Query Match	82.3%; Score 1185.8; DB 9; Length 1200;				
Best Local Similarity	99.8%; Pred. No. 0;				
Matches 1187; Conservative	0; Mismatches 2; Indels 0; Gaps 0;				
QY	211	GAGCTGGGAAATGGCCCGCAGCCGCGCGCTTCTCCTCTGATGCTTTTGCGAGA	270		
Db	1	GAGCTGGGAAATGGCCCGCAGCCGCGCGCTTCTCCTCTGATGCTTTTGCGAGA	60		
QY	271	CCGCGGTCTGCGAGGGCGCTGCTCGGTTTCACACTCTCTGCACCTTCATGGGTG	330		
Db	61	CCGCGGTCTGCGAGGGCGCTGCTCGGTTTCACACTCTCTGCACCTTCATGGGTG	120		
QY	331	CCTCAGCAGGACCTTGCTCTTCTCTTGAAGCTTTGGCTACGTGATGACCAGC	390		
Db	121	CCTCAGCAGGACCTTGCTCTTCTTGAAGCTTTGGCTACGTGATGACCAGC	180		
QY	391	TGTTCTGTTCTATGATGATGAGAGTGCCTGCTGAGCCCGCACTCCATGGTTTCCA	450		
Db	181	TGTTCTGTTCTATGATGATGAGAGTGCCTGCTGAGCCCGCACTCCATGGTTTCCA	240		
QY	451	GTAGAATTTCAACCCAGATGTGCTCAGCTGAGTGCAGAGTCTGAAAGGTGGATCACA	510		
Db	241	GTAGAATTTCAACCCAGATGTGCTCAGCTGAGTGCAGAGTCTGAAAGGTGGATCACA	300		
QY	511	TGTTCACTGTGACTTCTTGACTATTATGGAATCACAACACACAGAGGATCCCA	570		
Db	301	TGTTCACTGTGACTTCTTGACTATTATGGAATCACAACACACAGAGGATCCCA	360		
QY	571	CCCTGAGGTATCCTGGGCTGTTGAATGCAAGAACAACACAGTACCGAGGGCTACTGGA	630		
Db	361	CCCTGAGGTATCCTGGGCTGTTGAATGCAAGAACAACACAGTACCGAGGGCTACTGGA	420		
QY	631	AGTACGGGTATGATGGCAGGACACCTTGATTCCTGCTGACACACATGGATGGAGAG	690		
Db	421	AGTACGGGTATGATGGCAGGACACCTTGATTCCTGCTGACACACATGGATGGAGAG	480		
QY	691	CAGCAGAACCCAGGGCTGGCCCAACAAGCTGAGTGGGAAAGGCACAGATTCGGGCCA	750		
Db	481	CAGCAGAACCCAGGGCTGGCCCAACAAGCTGAGTGGGAAAGGCACAGATTCGGGCCA	540		
QY	751	GGCAGACAGGGCTACCTGGAGAGGAGTGCCTGCACAGCTGCAGCAGTTGCTGGAGC	810		
Db	541	GGCAGACAGGGCTACCTGGAGAGGAGTGCCTGCACAGCTGCAGCAGTTGCTGGAGC	600		
QY	811	TGGGGAGAGGTGTTTGGACCAAGTGCCTCTTGTGTGAAGGTGCACATCATGTGA	870		
Db	601	TGGGGAGAGGTGTTTGGACCAAGTGCCTCTTGTGTGAAGGTGCACATCATGTGA	560		
QY	871	CCTCTTCAGTGACACTCTACGTGTGCGGCCCTTGAACTACTACCCCAACAACATCACCA	930		
Db	661	CCTCTTCAGTGACACTCTACGTGTGCGGCCCTTGAACTACTACCCCAACAACATCACCA	720		
QY	931	TGAAGTGGCTGAAGGATACGACCCCAATGGATGCCAAGAGTTCGAACCTAAAGCGTAT	990		
Db	721	TGAAGTGGCTGAAGGATACGACCCCAATGGATGCCAAGAGTTCGAACCTAAAGCGTAT	780		
QY	991	TGCCCAATGGGATGGGACCTACAGGGCTGGGATACCTTGGCTGTACCCCTGGGGAAG	1050		
Db	781	TGCCCAATGGGATGGGACCTACAGGGCTGGGATACCTTGGCTGTACCCCTGGGGAAG	840		
QY	1051	AGCAGAGATATACGTACCAAGGTGGAGACCCAGGCTGGATCAGCCCTCATTTGATCT	1110		
Db	841	AGCAGAGATATACGTACCAAGGTGGAGACCCAGGCTGGATCAGCCCTCATTTGATCT	900		

QY	1111	GGGAGCCCTACCCTCTGGCACCCTAGTCATTTGAGTCAATCAGTGAATTCCTGTTTTG	1170		
Db	901	GGGAGCCCTACCCTCTGGCACCCTAGTCATTTGAGTCAATCAGTGAATTCCTGTTTTG	960		
QY	1171	TCGTCATCTTCTTTCATTCGAATTTTGTTCATAATTAAGCAAGGACGAGGTTCAAGAG	1230		
Db	961	TCGTCATCTTCTTTCATTCGAATTTTGTTCATAATTAAGCAAGGACGAGGTTCAAGAG	1020		
QY	1231	GAGCCATGGGCACCTACGCTTAGCTGAACCTGAGTGACACGACCTGCAGACTCACTG	1290		
Db	1021	GAGCCATGGGCACCTACGCTTAGCTGAACCTGAGTGACACGACCTGCAGACTCACTG	1080		
QY	1291	TGGGAAGGAGACAAAACCTAGAGACTCAAAGAGGAGTGCATTTATGAGCTTTCATGTTT	1350		
Db	1081	TGGGAAGGAGACAAAACCTAGAGACTCAAAGAGGAGTGCATTTATGAGCTTTCATGTTT	1140		
QY	1351	CAGGAGAGAGTTGACCTAAACATAGAAATTCCTGACGAACTCCTTCA	1399		
Db	1141	CAGGAGAGAGTTGACCTAAACATAGAAATTCCTGACGAACTCCTTCA	1189		
RESULT	12				
AF079407					
LOCUS	AF079407	1045 bp	mRNA	linear	PRI 18-MAR-1999
DEFINITION	Homo sapiens hemochromatosis splice variant dell14E4 (HFE) mRNA,				
complete cds.					
ACCESSION	AF079407				
VERSION	AF079407.1	GI:3695106			
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1045)				
AUTHORS	Rhodes,D.A. and Trowsdale,J.				
TITLE	Alternate splice variants of the hemochromatosis gene Hfe				
JOURNAL	Immunogenetics 49 (4), 357-359 (1999)				
MEDLINE	99180629				
REFERENCE	2 (bases 1 to 1045)				
AUTHORS	Rhodes,D.A.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-JUL-1998) Immunology, University of Cambridge, Tennis				
	Court Road, Cambridge CB2 1QP, UK				
FEATURES	Location/Qualifiers				
source	1..1045				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="6"				
	/map="6p22.1"				
gene	1..1045				
	/gene="HFE"				
CDS	37..1041				
	/gene="HFE"				
	/codon_start=1				
	/product="hemochromatosis splice variant dell14E4"				
	/protein_id="AAC62646.1"				
	/db_xref="GI:3695107"				
	/translation="MGPRAPALLMLLOTAVLQGRLLRSLHLYFMGASBQDLGL				
	SLEFALGYVDDQLFVFDHESRRVETPWSSRISSQMWLQLSQSLKGDHMTVDV				
	WTIMENHNHKSHTLQVILGCEMDENSTEGYWKYGYDGDHLEFCDPDLTLDWRAAEP				
	RAMPTKLEWERHRIARONRAYLERDPAQLLELGLGVLDQVPLVKTTHVTS				
	SVYTLRCALNYYPONITMKWLKQPMDAKEFEKDPNGDGYOGWITLAVPPGE				
	EORYTCOVHEPGLDPLVWEPSPSLVIGVISGIAVFVILFIFILRKRG				
	SRGANGHYVLAERE"				
	QNTIMKWLKQPMDAKEFEKDPNGDGYOGWITLAVPPGEORYTCOVHEPGLD				
	QPLIVIWEPSPSLVIGVISGIAVFVILFIFILRKRGSRGANGHYVLAERE				
BASE COUNT	243 a	259 c	314 g	229 t	
ORIGIN					
Query Match	68.7%; Score 989.8; DB 9; Length 1045;				
Best Local Similarity	96.0%; Pred. No. 7.7e-281;				
Matches 1043; Conservative	0; Mismatches 2; Indels 42; Gaps 1;				
QY	186	GAGATTTACGGGGACGTGCGCCACAGACTGGGGAATGGCCCCCGCAGACCCGCGC	245		

```
Db 1 GAGATTAAAGGGAGCTGCGCCAGAGCTGGGAAATGGGCCGCGAGCCAGCGCGCG 60
QY 246 CTTCTCCTCCTGATGCTTTTCAGACCGCGCTTCCTGCAGGGCGCTTGTGCGTTTCACAC 305
Db 61 CTTCTCCTCCTGATGCTTTTCAGACCGCGCTTCCTGCAGGGCGCTTGTGCGTTTCACAC 120
QY 306 TCCTCTCACTACCTCTTCATCGTGGTGCCTCAGACGAGGACCTTGGCTTTTCCTTGTGAA 365
Db 121 TCCTCTCACTACCTCTTCATCGTGGTGCCTCAGACGAGGACCTTGGCTTTTCCTTGTGAA 180
QY 366 GCTTTGGGTACGTGATGACACAGCTGTTCTGTTCTATGATGATGAGAGTCGCCGTGTG 425
Db 181 GCTTTGGGTACGTGATGACACAGCTGTTCTGTTCTATGATGATGATGATGATGATGATGAT 240
QY 426 GAGCCCGGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTTGGCTGCAGCTGAGT 485
Db 241 GAGCCCGGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTTGGCTGCAGCTGAGT 300
QY 486 CAGAGTCTGAAAGGGTGGATCACATGTTCTCACTGTTGACTTCTGGACTATTATGGAAT 545
Db 301 CAGAGTCTGAAAGGGTGGATCACATGTTCTCACTGTTGACTTCTGGACTATTATGGAAT 360
QY 546 CACAACCCACAGAGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAATGCAAGAA 605
Db 361 CACAACCCACAGAGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAATGCAAGAA 420
QY 606 GACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGCAGGACCCACCTTGAATTC 665
Db 421 GACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGCAGGACCCACCTTGAATTC 480
QY 666 TGCCTCTGACACACTGGATTGGAGAGCAGACAGCCAGGCGCTGGCCACCAAGCTGGAG 725
Db 481 TGCCTCTGACACACTGGATTGGAGAGCAGACAGCCAGGCGCTGGCCACCAAGCTGGAG 540
QY 726 TGGGAAAGCACAAGATTTCGGCCAGGCGAGCAGACAGGCGCTACTGGAGGAGCTGCCT 785
Db 541 TGGGAAAGCACAAGATTTCGGCCAGGCGAGCAGACAGGCGCTACTGGAGGAGCTGCCT 600
QY 786 GCACAGCTCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGCCTCT 845
Db 601 GCACAGCTCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACA----- 650
QY 846 TTGGTGAAGGTGACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGGGCGCTTG 905
Db 651 -----AGTGACCACTCTACGGTGTGGGCGCTTG 678
QY 906 AACTACTACCCCGAGACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGGATGCC 965
Db 679 AACTACTACCCCGAGACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGGATGCC 738
QY 966 AAGGAGTTTGAACCTAAAGACCTGATTGCCAATGGGATGGGACCTACAGGCGTGGATA 1025
Db 739 AAGGAGTTTGAACCTAAAGACCTGATTGCCAATGGGATGGGACCTACAGGCGTGGATA 798
QY 1026 ACCTTGGCTGTACCCCTCGGGGAGAGCAGAGATATACGTACCAAGTGGAGCACCAGGC 1085
Db 799 ACCTTGGCTGTACCCCTCGGGGAGAGCAGAGATATACGTGCGCAGTGGAGCACCAGGC 858
QY 1086 CTGGATCAGCCCTCATTTGTGATCTGGAGGCGCTCACGCTTGGGACCCCTAGTCAATTGA 1145
Db 859 CTGGATCAGCCCTCATTTGTGATCTGGAGGCGCTCACGCTTGGGACCCCTAGTCAATTGA 918
QY 1146 GTCATCAGTGAATTTGCTTTTCTGTCATCTTGTTCATTGGAATTTTGTTCATAATA 1205
Db 919 GTCATCAGTGAATTTGCTTTTCTGTCATCTTGTTCATTGGAATTTTGTTCATAATA 978
QY 1206 TTAAGGAAGAGGAGGGTTTCAAGAGGAGCCATGGGGCACTACGCTTTAGCTGAAGTGA 1265
Db 979 TTAAGGAAGAGGAGGGTTTCAAGAGGAGCCATGGGGCACTACGCTTTAGCTGAAGTGA 1038
QY 1266 TGACAGC 1272
|||||
```

```
Db 1039 TGACAGC 1045

RESULT 13
AY007541
LOCUS
DEFINITION Ceratotherium simum HFE protein mRNA, complete cds.
ACCESSION AY007541
VERSION AY007541.1 GI:10945687
KEYWORDS
SOURCE white rhinoceros.
ORGANISM Ceratotherium simum
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Ceratotherium.
REFERENCE 1 (bases 1 to 1320)
AUTHORS West,C.J., Worley,M. and Beutler,E.
TITLE Rhinoceros HFE Polymorphisms
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1320)
AUTHORS West,C.J., Worley,M. and Beutler,E.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Molecular and Experimental Medicine, The
Scripps Research Institute, 10550 North Torrey Pines Road, La
Jolla, CA 92037, USA
FEATURES
source Location/Qualifiers
1..1320
/organism="Ceratotherium simum"
/db_xref="taxon:9807"
86..1132
/codon_start=1
/product="HFE protein"
/protein_id="AAG23701.1"
/db_xref="GI:10945688"
/translation="MGPRARPALFFLLRTVAAGRPSPRSHSLRYLPMGASERDHL
PLEALGYVDDLEFVAINESRRAESRAQNVLGEARSLWLQSLKSGNDHMFIVDF
RLTMDNNHNSKESHTIQVILGCEVEDNSTRGFWKYGDQDHFCEPETLDWRAAS
RAITTKLEWENKIRAKONRAYLRDCPEQLWLLLEGLVDQDQVPLVYKTHHVAS
AVTTLRQALNFYQNTWRWLKRPMDVKDAESKDVLPSDGTQSWEALAVPPE
EORTTCQVEHPGLDPLATWEPSTLNTVTGTVISGIAVCVITFIIFIGILFRLKRQA
SRGAMGDYVLGECE"
BASE COUNT 284 a 363 c 394 g 279 t
ORIGIN

Query Match 62.1%; Score 893.8; DB 4; Length 1320;
Best Local Similarity 82.2%; Pred. No. 1.8e+252;
Matches 1063; Conservative 0; Mismatches 227; Indels 3; Gaps 3;

QY 137 AGGTGACTTCTGAGGCCATCCCGGTTTCCCGCCCCCAAGAGCGGAGATTTAACG 196
Db 2 AGGTGACTTCTTGGGTCACTCTCTATTTCCCGCCCCCAGCAAGCAGCTGA-CGCGGGCC 60

QY 197 GGGACGTGCGGCCAGACTGGGNAATGGCCCGCGAGCCGAGCCGCGCTTCTCTCCT 256
Db 61 AGCACCGCGGTCCGAGCGGGGAAATGGCCCGCGAGCCGCGCTCTTCTTCTCT 120

QY 257 GATCCTTTTCAGACCGCGCTCTGACGGCGCTTCTCGCTTTCACACTCTCTGCACCTA 316
Db 121 GATCCTCTCGGACCGCTGGCCGGCAGGGGCGACCAACCGCGGTCTCTCTCGGCTA 180

QY 317 CCTTTCATGGTGCCCTCAGAGCAGGACCTTGGTCTTCTTCTTCTTCTTGAAGCTTTGGGCTA 376
Db 181 CCTTTCATGGTGCCCTCAGAGCGGGACCAAGCGCTTCCCTTCTTGAAGCTTTGGGCTA 240

QY 377 CGTGGATGACCAGCTGTCTGTTCTATGATGATGAGAGTCGCGCTGTGGAGCCCCGAAC 436
Db 241 TGTGGATGACGAGCTTCTGTTGGCTTACAAATCAGAGAGTCGCCGTGCGAGAGTCCCCTGC 300

QY 437 TCCATGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA 496
Db 301 CCAGTGGGTCTTGGGTGAGGCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCTGAA 360

QY 497 AGGTGGGATCACATGTTTCACTGTGACTTCTTGACTTATTTGAAATATCAACACACAG 556
```


[illegible]

RESULT	14
AY007543	
LOCUS	AY007543
DEFINITION	Dicerorhinus sumatrensis HFE protein mRNA, complete cds.
ACCESSION	AY007543
VERSION	AY007543.1 GI:10945691
KEYWORDS	.
SOURCE	Sumatran rhinoceros.
ORGANISM	Dicerorhinus sumatrensis
	linear MAM 22-OCT-2000

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Dicerorhinus. 1 (bases 1 to 1320)
 AUTHORS West,C.J., Worley,M. and Beutler,E.
 TITLE Rhinoceros HFE Polymorphisms
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1320)
 AUTHORS West,C.J., Worley,M. and Beutler,E.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) Molecular and Experimental Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
 source Location/Qualifiers
 1..1320
 /organism="Dicerorhinus sumatrensis"
 /db_xref="taxon:89632"
 86..1132
 /codon_start=1
 /product="HFE protein"
 /protein_id="AAG23703.1"
 /db_xref="GI:10945692"
 /translation="MGPRARPALFFLLILRTVAAGPPRSHSLRYLFMGASERDGLPLPFLPAGVDELDFVAYNSRAESRAQWLVGEAHSQLWLQLSQSLKGDHMFIVDFWTMDNHNHKSSELTLOVILGCEVOENSTRGFKYGDGDHLEFCPETLDMPRAASRALTTKLEWENVKIRAKONRAYLERDCPEQLWILLELRGLVLDQVPLVKVTHHVASAVTTTCQALNFYPNITMRLKVDKDAESKDVLPFGSDGTYSOSWALAVPGEAEVRYTCQVHPGLDPLTATWEPSSLNTLVGTVISIAVCVIFILIGILFRILKRKQA SRGAMGDYLAECB"

BASE COUNT 281 a 358 c 398 g 283 t

ORIGIN

	Query Match	62.1%;	Score 893.8;	DB 4;	Length 1320;
	Best Local Similarity	82.2%;	Pred. No. 1.8e-252;		
	Matches 1063;	Conservative 0;	Mismatches 227;	Indels 3;	Gaps 3;
QY 137	AGGTGACTTCTGGAGCATCCCGTTCCCGCGCCCCCAAGAGCGGAGATTAAACG	196			
Db	2 AGGTGACTTCTGGGTCACTCTATTTCGCGCCCAACGACAGCTGACA-CGGGCC	60			
QY 197	GGAGCTGGCGCAGAGCTGGGAAATGGGCCGGGAGCGCGGCGCTTCTCCTCT	256			
Db	61 AGCACCGGGTCCGAGCGGGGAATCGGCCGGGAGCCGACCCGCGCTTCTTCT	120			
QY 257	GATGCTTTGAGACCGCGGTCCTGACGGGCGCTTGCTGGCTTACACTCTCTGCACTA	316			
Db	121 GATCTCTCGGAACCGTGGCGCGCAGGGCGACCGCGGGTCACTCTCTGCGCTA	180			
QY 317	CCTCTCATGGGTGCCTCAGACAGGACCTTGGCTTTCTTGTGTTGAAGCTTTGGGCTA	376			
Db	181 CCTCTCATGGCGCCTCAGACGGGACACGGCGCTGCCCTGTTTGAAGCTTTGGGCTA	240			
QY 377	CGTGATCACCAGCTGTTCGTGTTCTATGATGATGAGAGTCGCGTGTGGAGCCCCGAAC	436			
Db	241 TGTGATCAGCAGCTTTTCGTGGCGTCAATCATGAGAGTCGCGTGCAGAGTCCCCTGC	300			
QY 437	TCCATGGTTCCTCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAGAGTCTGAA	496			
Db	301 CCAGTGGGTCTTGGGTGAGGCCACAGCCAGCTGTGGCTGCAGCTGAGTCAGAGCTGAA	360			
QY 497	AGGTGGGATCACATGTTCTACTGTGACTCTGGACTATTATGGAATAATCACACACAG	556			
Db	361 AGGTGGGATCACATGTTCTATCGTCGACTTCTGGACCATCATGGACAACCAACACAG	420			
QY 557	CAAGAGTCCACACCCCTGCAGGTCATCTGGGCTGTCAATATGAAGNAGACACAGTAC	616			
Db	421 CAAGAGTCCACACCCCTGCAGGTCATCTGGGCTGTGAGGTGCAAGAGGACAACAGCA	480			
QY 617	CGAGGGCTACTTGGAGTACGGGTATGTGGCAGGACCACTTGAATTTCTGCCCTGACAC	676			
Db	481 CAGGGGTTCTTGAAGTACGGGTACGATGGCAGGACCATCTTGAATTTCTGCCCTGAGAC	540			
QY 677	ACTGATTTGAGAGCAGCAGAACCCAGGGCCTGGCCCCACCAAGCTGGAGTGGGAAAGCA	736			

|||||
Db 541 ACTGGATTGGAGAGCAGAGTCCAGGCGCTGACCCACCAAGCTGGAGTGGAAAGTGAA 600
QY 737 CAAAGATCGGGCCAGCAGCAAGAGGCGCTACCTGGAGAGGACTGCCCTGCACAGCTGCA 796
Db 601 CAAAGATCGGGCCAGCAGCAAGAGGCGCTATCTCGAGAGAGACTGCCCTGAGCAGCTGCA 660
QY 797 GCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGCCTCTCTTTGGTCAAGGT 856
Db 661 GTGTTGCTGGAGCTGGGAGAGGTTCTGGACCAAGTGCCTCTCTTTGGTCAAGGT 720
QY 857 GACACATATGTAGCTCTTCAGTGACCACTCTACGGGTGTCGGGCGCTTGAACCTACTACCC 916
Db 721 GACTCATCATGTGGCGCTCTGCAGTGACCACTCTACGGGTGTCAGGCTCTGAACCTTCTACCC 780
QY 917 CCAGACATCACCATCAAGTGCCTGAAGGATAGCAGCAATGGATGCCAGGAGTTTGA 976
Db 781 CCAGAAATCAGCATGAGTGGTTGAAGACAGGAAGCAGTGGATGTCAAGGACGCCGA 840
QY 977 ACTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCGCTGGATAACCTTGGCTGT 1036
Db 841 GTCTAAGGATGTGCTGCCAGCGGGATGGGACCTACCAAGACTGGTGGCTTTGGCGGT 900
QY 1037 ACCCGCTGGGGAAGACAGAGATATACGTACCAGGTGGAGCAACCCAGGCGCTGGATCAGCC 1096
Db 901 GCCCGCTGGGGAAGACAGAGATATACCTGCCAGGTGGAGCAACCCAGGCGCTGGATCAGCC 960
QY 1097 CCTCATTTGTATCTGGAGCGCTCACCCTGCGGACCCCTAGTCACTTGGAGTCAATCAGTGG 1156
Db 961 CCTCACTGCCACTGGGAGCGCTCAGCTTCTAACACCTTGTCACTGGAGTCAATCAGTGG 1020
QY 1157 AATTGCTGTTTTGTGCTCATCTTGTTCATTGGAATTTTGTTCATAATTAAGGAAGAG 1216
Db 1021 GATTGCTGTTTGTGTCATCATCTTCTCATTTGGAATTTTGTTCAGAATCTTAAGGAAAAG 1080
QY 1217 GCAGGTTTCAAGAGAGCCATGGGCGACTACGCTCTAGCTGAACCTGAGTGACAGCAGC 1276
Db 1081 ACAAGCTTCGAGAGGAGCCATGGGAGCTACGCTCTAGCTGAATCTGAGTGAAGCAGCAGC 1140
QY 1277 CTGCAGACTCACTGTGGGAAGGAGCAAACTAGAGACTCAAGAGGAGGTGCATTTATG 1336
Db 1141 CTGCAGACTCCTGTGGGAGGAGCAGACATAATTGGAGACTCAGAGAGGTGCAC-TTGTGG 1199
QY 1337 AGCTCTTATGTTTCAGGAGAGAGTTGAACCTAA-ACATAGAAATTGCCTGACGAACCTCC 1395
Db 1200 CECTTCTCTGTTTTCAGGACAGAGTTGAACCTTAACAAATGGGAAGTGTCCCGAGGAGTCT 1259
QY 1396 TTGATTTAGCCTCTCTCTTCACTTCCCTCAA 1428
Db 1260 CTGCTCTTAGCCTTCCCTGTTCACTCTCAAAA 1292

RESULT 15
AY007544
LOCUS AY007544 1319 bp mRNA linear MAM 22-OCT-2000
DEFINITION Rhinoceros unicornis HFE protein mRNA, complete cds.
ACCESSION AY007544
VERSION AY007544.1 GI:10945693
KEYWORDS greater Indian rhinoceros.
SOURCE Rhinoceros unicornis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Rhinoceros.
1 (bases 1 to 1319)
TITLE Rhinoceros HFE Polymorphisms
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1319)
AUTHORS West,C.J., Worley,M. and Beutler,E.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Molecular and Experimental Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
source Location/Qualifiers
1..1319
/organism="Rhinoceros unicornis"
/db_xref="taxon:9809"
86..1132
/codon_start=1
/product="HFE protein"
/protein_id="AAG23704.1"
/db_xref="GI:10945694"
/translation="MGPRAPALFFLLILRTVAAGRRPRSHSLRYLPMGASERDHL
PLEALGYVDELFFVAINHESRAESRAQWVLEAHSQWLQSLGSKGWDHMFIVDF
WTMDNNHNSKESHTLQVILGCEVEDNSTRGFWMYGYDQDHLSEFCPETLWRAAS
RALTTKLEWENKIRAKONRLEYLERCPEQLWLLLELGRVLDDQVPLVKYTHVAS
AVTTLKQALNFPQNTMRWLKDRKPDVDAESKDLVPSGDGTQSVWALVPPGE
EORYTCQVEHPGLDPLTATWEPSSLNTLVGTVISGIAVCVIFILIGILFRLRKROA
SRGAGDYVLAEC"
BASE COUNT 279 a 363 c 397 g 280 t
ORIGIN
Query Match 62.1%; Score 893.6; DB 4; Length 1319;
Best Local Similarity 81.7%; Pred. No. 2.1e-252;
Matches 1056; Conservative 0; Mismatches 234; Indels 2; Gaps 2;
QY 137 AGGTGACTTCTGGAGCCATCCCGGTTTCCCGCCCCCAAGAGCGAGATTTAACG 196
Db 2 AGGTGACTTCTTGGTCACTCTCTATTTCCCGCCCCCAAGAGCGAGCTGACA-CGGGCC 60
QY 197 GGGAGCTGCGGCCAGAGCTGGGNAATGGCCCCCGAGCGCGCGCTTCTCTCCT 256
Db 61 AGACCCCGCGGTCCGAGCGGGGAAATGGCGCGGAGCGCGCGCTCTCTTCTCT 120
QY 257 GATCTTTTTCAGACACCCTGCTGTCAGGGCGCTTCTGCTTTCACACTCTCTGCACTA 316
Db 121 GATCTCTCTGGACCCGTGGCGGCGAGGGCGACCCCGGCTCACACTCTCTGCGCTA 180
QY 317 CCTTTCATGGTGCCCTCAGAGAGGACTTGGTCTTCTTCTTCTTGTGAAGCTTTGGGCTA 376
Db 181 CCTTTCATGGTGCCCTCAGAGCGGACCTAGCTGCCCTGCCCTGTTTGAAGCTTTGGGCTA 240
QY 377 CGTGATGACAGCTGTCTGTTCTATGATGAGTGCAGTGCCTGTGAGCGCCGAAAC 436
Db 241 TGTGATGATGAGCTTTTCTGGCTACATACAGAGAGTGCCTGTGAGAGTCCCGTGC 300
QY 437 TCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGCTGCAGCTGAGTCAGAGTCTGAA 496
Db 301 CCAGTGGGTCTTGGGTGAGGCCACAGCCAGCTGTGCTGCAGCTGAGTCAGAGCTGAA 360
QY 497 AGGTTGGATCACATGTTCACTGTTGACTTCTGACTATTATGAAATCACACACCAG 556
Db 361 AGGTTGGATCACATGTTCACTGTTGACTTCTGACTATTATGAAATCACACACCAG 420
QY 557 CAAGGAGTCCACACACCTCGAGGCTCATCTGGGCTGTGAATGCAAGAAGCAACAGTAC 616
Db 421 CAAGGAGTCCACACACCTCGAGGCTCATCTGGGCTGTGAGTGCAGAGGACACAGCAC 480
QY 617 CGAGGGTACTTGAAGTACGGGTATGATGGGAGGAGCAACCTTTGAATCTGCCCTGACAC 676
Db 481 CAGGGGTCTTGGATGTACGGGTACGAGCGGAGGAGCACTTCTGAATCTGCCCTGAGAC 540
QY 677 ACTGATTTGAGAGCAGCAGACCCAGGCGCTGGCCCAACCAAGCTGGAGTGGGAAGGCA 736
Db 541 ACTGATTTGAGAGCAGCAGAGTCCAGGGCGCTGACCAACCAAGCTGGAGTGGGAAGTGA 600
QY 737 CAAGATTCGGGCCAGCAGCAAGAGGCGCTACCTGGAGAGGAGTGCCTGCACAGCTGCA 796
Db 601 CAAGATTCGGGCCAAGCAGACAGAGCGCTATCTCGAGAGGAGTGCCTGCAGAGCTGCA 660
QY 797 GCAGTTGCTGGAGCTGGGAGAGAGTGTGTTGGACCAACAAGTGCCTCTTGGTGAAGT 856
Db 661 GTGTTGCTGGAGCTGGGAGAGGAGTGTGTTGGACCAACAAGTGCCTCTTGGTGAAGT 720
QY 857 GACACATCATGTGACCTTTCAGTGACCACTCTACGGGTGTCGGGCGCTTGAACCTACTACCC 916

Db 721 GACTCATATGTGGCTCTGCAAGTACCACCTCTACGGTGTAGGCTCTGAACCTTCTACCC 780
QY 917 CCAGAAACATCACCATGAAGTGAAGGATAAGCAGCAATGGATGCCAAGGAGTTGGA 976
Db 781 CCAGAAACATCACCATGAGTGGTTGAAGGACAGGAGCAGTGGATGTCAAGGACGCCGA 840
QY 977 ACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAAGGCTGGATAACCTTGGCTGT 1036
Db 841 GTCTAAGGATGTGCTGCCAGCGGGGATGGGACCTACCAGAGCTGGGTGGCTTTGGCTGT 900
QY 1037 ACCCCCTGGGGAAGACAGACATATACGTACCAGGTGGAGCACCAGCCCTGGATCAGCC 1096
Db 901 GCCCCCCGGGGAAGACAGAGATATACCTGCCAGGTGGAGCACCAGGCCCTGGATCAGCC 960
QY 1097 CTTCAATGTGATCTGGGAGCCCTACCCGCTCGGCACCCCTAGTCAATGGAGTCATCAGTGG 1156
Db 961 CCTCACTGCCACCTGGGAGCCCTCACTTTCTAACACCCCTCGTCACTGGAGTCATCAGTGG 1020
QY 1157 AATTGCTGTTTTTTCGTCATCTTGTTCATTTGGAATTTTGTTCATAATTAAGGAAGAG 1216
Db 1021 GATTGCTGTTTGTGTCATCTCTCTCATTTGGAATTTTGTTCAGAACTTTAAGGAAAAG 1080
QY 1217 GCAGGGTTCAAGAGGAGCCATGGGCACTTACGCTTTAGCTGAACGTGAGTGACACGCAGC 1276
Db 1081 ACAAGCTTCAAGAGGAGCCATGGGGACTACGCTTACGTGAATGTGAGTGAAGCACAGC 1140
QY 1277 CTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAAAGAGGAGTGCAATTTATG 1336
Db 1141 CTGCAGACTCCTTGTGGGAGGACACATAATTTGGAGACTCAGAGAGGGTGCAC-TGGTGG 1199
QY 1337 AGCTCTTCATGTTTCAGGAGAGAGTTGACCTAAACNTAGAAAATTCCTGACGAACCTCT 1396
Db 1200 CGTTCTCTGTTTCAGGACAGAGTTGAACCTTAACAAATGGGAAGTCCAGGAGGTCT 1259
QY 1397 TGATTTTAGCCTTCTCTGTTTCATTCTCAAA 1428
Db 1260 CTCTCTAGCCTTCCCTGTTCACTCTCAAAA 1291

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 07:21:40 ; Search time 2546.54 Seconds
(without alignments)
7632.166 Million cell updates/sec

Title: US-09-497-957-12
Perfect score: 1440
Sequence: 1 GGGGACTGGATCACCTAG.....TCCTCAAAAGATTCCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estinu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_esti.*
10: gb_est2.*
11: gb_hic.*
12: gb_gss.*
13: em_gss_hum.*
14: em_gss_inv.*
15: em_gss_pln.*
16: em_gss_vrt.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Query Match	Length	ID		
1	735.4	51.1	819	10	BG747345	BG747345 602704818
2	630.6	43.8	1723	11	AK009581	AK009581 Mus muscu
3	505.2	35.1	570	10	BE272926	BE272926 601171213
c 4	356.4	24.8	384	10	BF883952	BF883952 PM4-ET020
5	282.6	19.6	550	10	BI339179	BI339179 364041 MA
6	275	19.1	523	10	BF080089	BF080089 230846 MA
7	257.8	17.9	489	10	BE94943	BE94943 UI-M-CG0p
8	251.4	17.5	457	9	A1850020	A1850020 UI-M-BG0-
9	248.4	17.2	455	10	BE995172	BE995172 UI-M-CG0p
c 10	230	16.0	464	9	AA217236	AA217236 mu89b05.r
11	226.4	15.7	268	10	W21141	W21141 zb51g06.r1
12	214.6	14.9	546	10	BE809138	BE809138 214520 MA
13	209.6	14.6	481	9	BB851691	BB851691 BB851691
14	202.2	14.0	831	10	BI452668	BI452668 603169877
15	197.2	13.7	502	9	BB858165	BB858165 BB858165
16	196.6	13.7	392	10	BF465475	BF465475 UI-M-CG0p
17	184.6	12.8	1208	11	BC005306	BC005306 Homo sapi

18	174	12.1	394	10	BF464345	BF464345 UI-M-CG0p
c 19	168.8	11.7	444	12	AZ025590	AZ025590 RPCI-23-3
c 20	160	11.1	334	9	AW902003	AW902003 QVO-NN102
c 21	155.2	10.8	536	12	AZ074871	AZ074871 RPCI-23-4
22	154	10.7	805	10	BG402460	BG402460 602466163
23	152	10.6	1016	9	AL532717	AL532717 AL532717
24	151.8	10.5	289	10	H33644	H33644 EST109630 R
c 25	150	10.4	481	12	AZ025784	AZ025784 RPCI-23-3
26	146	10.1	531	12	AQ703007	AQ703007 HS_5443_B
27	144	10.0	1031	9	AL513728	AL513728 AL513728
28	142	9.9	1298	11	AK005051	AK005051 Mus muscu
29	141.6	9.8	1121	9	AL545738	AL545738 AL545738
30	140	9.7	972	9	AL541209	AL541209 AL541209
31	139.8	9.7	959	9	AL560013	AL560013 AL560013
32	138.2	9.6	910	9	AL558822	AL558822 AL558822
33	137.8	9.6	1020	9	AL535151	AL535151 AL535151
34	137.6	9.6	942	9	AL539924	AL539924 AL539924
35	136.4	9.5	1094	10	BM465674	BM465674 AGENCOURT
36	132.6	9.2	932	9	AL547313	AL547313 AL547313
37	131.8	9.2	679	10	BG680802	BG680802 602628602
38	131.4	9.1	538	9	BE232472	BE232472 137492 MA
39	131	9.1	936	10	BI832514	BI832514 603082113
40	131	9.1	937	9	AL540488	AL540488 AL540488
41	131	9.1	1090	10	BM477276	BM477276 AGENCOURT
42	130	9.0	604	10	BG691825	BG691825 341215 BA
43	130	9.0	930	9	AL541347	AL541347 AL541347
44	130	9.0	979	9	AL542403	AL542403 AL542403
45	130	9.0	979	9	AL557373	AL557373 AL557373

ALIGNMENTS

RESULT 1
BG747345 819 bp mRNA linear EST 15-MAY-2001
LOCUS 602704818F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4857941 5',
DEFINITION mRNA sequence.
ACCESSION BG747345
VERSION BG747345.1 GI:14057998
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 819)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LINC1711 row: d column: 06
High quality sequence stop: 792.

FEATURES
source
1..819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4857941"
/clone.lib="NIH_MGC_15"
/tissue type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pORF7; Site:1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in

						the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)*						
BASE COUNT	202 a	201 c	235 g	181 t								
ORIGIN												
Query Match	51.1%;	Score	735.4;	DB 10;	Length	819;						
Best Local Similarity	95.9%;	Pred.	No. 3.8e-194;									
Matches	787;	Conservative	0;	Mismatches	31;	Indels	3;	Gaps				
Qy	384	GACCAGCTGTTCCGTTCATGATGATGAGAGTGC	CGCGTGTTGGAGCCCCGAAC	TCCATGG	443							
Dd	1	GACCAGCTGTTCCGTTCATGATCATGAGASTGC	CGCGTGTTGGAGCCCCGAAC	TCCATGG	60							
Qy	444	GTTTTCCAGTAGAATTCCAAGCCAGATG	GGCTGCAGCTGAGTCTCAAAAGG	TGG	503							
Dd	61	GTTTTCCAGTAGAANTTCAGGCCAGATG	GGCTGCAGCTGAGTCTGAAAAGG	TGG	120							
Qy	504	GATCACATGTTCACTGTTGACTTCTGGACTATT	TATGAAAAATCAACACAGCAAGG	GAG	563							
Dd	121	GATCACATGTTCACTGTTGACTTCTGGACTATT	TATGAAAAATCAACACAGCAAGG	GAG	180							
Qy	564	TCCCACACCTTGCAGGTCATCCTGGCTGTCAA	TATGCNAGAAGACAAGTACCGAGGGC	623								
Dd	181	TCCCACACCTTGCAGGTCATCCTGGCTGTGAAA	TATGCAAGAAGACAACAGTACCGAGGGC	240								
Qy	624	TACTGGAAGTACGGGTATGATGGCAGGACCACT	TGAATTTCTGCCCTGCACACTGGAT	683								
Dd	241	TACTGGAAGTACGGGTATGATGGCAGGACCACT	TGAATTTCTGCCCTGCACACTGGAT	300								
Qy	684	TGGAGACGACGAGAACCCAGGCGCTTGGCCCC	ACCACCAAGCTGGAGTGGGAAAGCACAAGATT	743								
Dd	301	TGGAGACGACGAGAACCCA - GGCTTGGCCCC	ACCACCAAGCTGGAGTGGGAAAGCACAAGATT	359								
Qy	744	CGGGCCAGGACAGACAGGGCGTACTTGGAGAG	GAGACTGCCCTGCACAGCTCCACGAGTTG	803								
Dd	360	CGGGCCAGGACAGACAGGGCGTACTTGGAGAG	GAGACTGCCCTGCACAGCTCCACGAGTTG	419								
Qy	804	CTGGAGCTGGGGAGAGTGTTTGGACCAACAAG	TGCCTCTTTGGTGAAGGTGACACAT	863								
Dd	420	CTGGAGCTGGGGAGAGTGTTTGGACCAACAAG	TGCCTCTTTGGTGAAGGTGACACAT	479								
Qy	864	CATGTGACCTCTTCAGTGCACACACTTACGGT	STGCGGCGTTTGAAC	TACTACCCCGCAGAAC	923							
Dd	480	CATGTGACCTCTTCAGTGCACACACTTACGGT	STGCGGCGTTTGAAC	TACTACCCCGCAGAAC	539							
Qy	924	ATCACCATGAAGTGGCTGAAGGATAAGCAGCA	ATGATGCCAAGGAGTTTCCGAACCTAAA	983								
Dd	540	ATCACCATGAAGTGGCTGAAGGATAAGCAG - CA	ATGATGCCAAGGAGTTTCCGAACCTAAA	598								
Qy	984	GACGTATTTGCCAATGGGATGGGACCTACCAG	GGCTGGATAACCTTGGCTGTACCCCT	1043								
Dd	599	GACGTATTTGCCAATGGGATGGGACCTACCAG	GGCTGGATAACCTTGGCTGTACCCCT	657								
Qy	1044	GGGGAAGACGACAGATATACATGACAGTG	GAGACCCAGGCGCTGGATCAGCCCTCATTT	1103								
Dd	658	GGGGAAGACGACAGATATACATGACAGTG	GAGACCCAGGCGCTGGATCAGCCCTCATTT	717								
Qy	1104	GTGATCTGGGAGCCCTCACCGCTCTGGCAC	CCCTAGTCAATGGAGTCAATGAGTGGATTC	1163								
Dd	718	GTGATCTGGGAGCCCTCACCGCTCTGGCAC	CCCTAGTCAATGGAGTCAATGAGTGGATTC	777								
Qy	1164	GTTTTTGTGCGTCATCTGTTTCATTTGGAAT	TTTGTTCATAAT	1204								
Dd	778	TGTTTTGTGCGTCATCTGTTTCATTTGGAAT	TTTGTTCATAAT	818								
RESULT 2												
AK009581						1723 bp	mRNA	linear	HTC 19-JAN-2002			
LOCUS												
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:231003ZM04:hemochromatosis, full insert sequence.											

ACCESSION	AK009581.1	GI:12844462
VERSION	HTC; CAP trapper.	
KEYWORDS	Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,	
SOURCE	clone_lib:RIKEN full-length enriched mouse cDNA library clone:2310032M04.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;	
AUTHORS	Carninci,P. and Hayashizaki,Y.	
TITLE	High-efficiency full-length cDNA cloning	
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)	
MEDLINE	99279253	
PUBMED	10349636	
REFERENCE	2 (sites)	
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE	20499374	
PUBMED	11042159	
REFERENCE	3 (sites)	
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawaji,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kirai,A. and Hayashizaki,Y.	
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE	20530913	
PUBMED	11076861	
REFERENCE	4 (sites)	
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
TITLE	Functional annotation of a full-length mouse cDNA collection	
JOURNAL	Nature 409, 685-690 (2001)	
REFERENCE	5 (bases 1 to 1723)	
AUTHORS	Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Horii,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Koike,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J., Saito,H., Saigo,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toyota,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ , Tel:81-45-503-9222, Fax:81-45-503-9216)	
COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGAGAGGATCGAAGCTCTTTTTTTTTTTTNN 3'. cDNA was prepared by using trihalose thermo-activated reverse transcriptase	

ACCESSION BE272926
VERSION BE272926.1 GI:9147279
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 570)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW240 row: j column: 04
High quality sequence stop: 566.
FEATURES
Location/Qualifiers
source
1..570
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3544803"
/clone.lib="NIH_MGC.14"
/tissue.type="renal cell adenocarcinoma"
/lab.host="DH10B (phage-resistant)"
/note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 140 a 148 c 175 g 107 t
ORIGIN
Query Match 35.1%; Score 505.2; DB 10; Length 570;
Best Local Similarity 99.2%; Pred. No. 6.3e-130;
Matches 518; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 560 GGAGTCCACACCCCTGCAGTGCATCTGGCTGTGAATGCAAGAGACACAGTACCGA 619
DB 48 GCAGTCCACACCGTGCAGTGCATCTGGCTGTGAATGCAAGAGACACAGTACCGA 107
QY 620 GGCCTACTGGAAGTACGGGTATGTGGCAGGACACACCTTGAAATCTGCCCTGACACACT 679
DB 108 GGCCTACTGGAAGTACGGGTATGTGGCAGGACACACCTTGAAATCTGCCCTGACACACT 167
QY 680 GGATTGGAGCAGCAGACACCCAG-GGCCTGCCACCAAGCTGCAGTGGGAAGGCACA 738
DB 168 GGATTGGAGCAGCAGACACCCAGTGGCTGCCACCAAGCTGCAGTGGGAAGGCACA 227
QY 739 AGATTGGGCCCGCCAGCAGACAGGGCCTACCTGGAGAGGAGTGCCTGCACAGCTGCAGC 798
DB 228 AGATTGGGCCCGCCAGCAGACAGGGCCTACCTGGAGAGGAGTGCCTGCACAGCTGCAGC 287
QY 799 AGTTGCTGAGCTGGGGAGAGGTGTTTGGACCAACAGTGCCTCTTTGGTGAAGGTGA 858
DB 288 AGTTGCTGAGCTGGGGAGAGGTGTTTGGACCAACAGTGCCTCTTTGGTGAAGGTGA 347
QY 859 CACATCATGTGACCTCTTCAGTGACCACCTCTACGGTGTGCGGCCCTTGAACCTACACCCC 918
DB 348 CACATCATGTGACCTCTTCAGTGACCACCTCTACGGTGTGCGGCCCTTGAACCTACACCCC 407
QY 919 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGGATGCCAAGGAGTTTGAAC 978
DB* 408 AGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCCAATGGATGCCAAGGAGTTTGAAC 467

QY 979 CTAAGACGCTATTGCCCAATGGGGATGGGACCTACCAGGCTGGATAACCTTGGCTGTAC 1038
DB 468 CTAAGACGCTATTGCCCAATGGGGATGGGACCTACCAGGCTGGATAACCTTGGCTGTAC 527
QY 1039 CCCTGGGGAGAGCAGAGATATACGTACCAGGTGGAGCAC 1080
DB 528 CCCTGGGGAGAGCAGAGATATACGTGGCAGGTGGAGCAC 569
RESULT 4
BF883952/c
LOCUS BF883952 384 bp mRNA linear EST 17-JAN-2001
DEFINITION PM4-ET0209-151200-003-f07 ET0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF883952
VERSION BF883952.1 GI:12274078
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 384)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-ET0209-
151200-003-f07&t3=2000-12-15&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 17
High quality sequence stop: 384.
FEATURES
Location/Qualifiers
source
1..384
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone.lib="ET0209"
/dev_stage="Adult"
/note="Organ: lung_tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No.196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 92 a 112 c 87 g 93 t
ORIGIN
Query Match 24.8%; Score 356.4; DB 10; Length 384;
Best Local Similarity 99.5%; Pred. No. 1.9e-88;
Matches 368; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 325 TGGCTGCCTCAGACAGACCTTGGTCTTCCCTTGTGTTTGAAGCTTTGGCTACGTGGATG 384
DB 384 TGGGTGCCTCAGACAGACCTTGGTCTTCCCTTGTGTTTGAAGCTTTGGCTACGTGGATG 325
QY 385 ACCAGCTGTTCTGTTCTATGATGATGAGTGCCTGTGGAGCCCCCAACTCCATGGG 444
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||


```

Db 324 ACCAGCTG-TCGTGTTCTATGATCATGAGAGTCGCCCTGTGGAGCCCCCGAACTCCATGGG 266
QY 445 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCTGAGAGTCTGAAAGGTTGGG 504
Db 265 TTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCTGAGAGTCTGAAAGGTTGGG 206
QY 505 ATCATGTTTCACATGTTGACTTCTGGACTATTATGNAATATCACACCACAGCAAGAGT 564
Db 205 ATCATGTTTCACATGTTGACTTCTGGACTATTATGNAATATCACACCACAGCAAGAGT 146
QY 565 CCCACACCTTCGAGGTCATCTCGGCTGTGAAATGCAAGAACACACAGTACCGAGGGCT 624
Db 145 CCCACACCTTCGAGGTCATCTCGGCTGTGAAATGCAAGAACACACAGTACCGAGGGCT 86
QY 625 ACTGGAAGTACGGGTATGATGGCAGGACCACTTGAATCTCCCTGACACACTGGATT 684
Db 85 ACTGGAAGTACGGGTATGATGGCAGGACCACTTGAATCTCCCTGACACACTGGATT 26
QY 685 GGAGAGCAGC 694
Db 25 GGAGAGCAGC 16

RESULT 5
BI339179
LOCUS 364041 MARC 2PTG Sus scrofa cDNA 5', mRNA linear EST 30-JUL-2001
DEFINITION 550 bp mRNA
ACCESSION BI339179
VERSION BI339179.1 GI:15032462
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 550)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGAGG
Plate: 100 row: C column: 24
Seq primer: ATTTAGGTGACACTATAG.
FEATURES
source
1..550
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PTG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 108 a 180 c 164 g 98 t
ORIGIN

Query Match 19.6%; Score 282.6; DB 10; Length 550;
Best Local Similarity 74.0%; Pred. No. 9.3e-68;
Matches 412; Conservative 0; Mismatches 114; Indels 31; Gaps 3;

```

```

QY 140 GTGACTTCTGAGCATCCCGTTTCCCGCCGCCCAAAAGAGCGGAGATTTAACGGGG 199
Db 1 GTGACTTCTGAGCCCTCGGTTTCCCGCCGCCCAACAGCGCCGAAAG-----CCCTGG 55
QY 200 AGCTGGGCGCAGAGTGGGAAATGGCCCGGAGCCAGCGCGGCGTCTCTCTCTGAT 259
Db 56 AAGCGCGTCCAGCGCGG--AAAGGCGCAACAAGCCCGCGGCGTCTCTCTCTGAT 113
QY 260 GCTTTTGCAGACCGGCTCCTCGAGGGGCTTGTGGTTTCACTCTCTGCACTACT 319
Db 114 CTTCTCGGACCGTGGCCACGAGGAGCGCGCGGCGCCACACTCCCTGCTTCTCT 173
QY 320 CTTATGGGTGCTTCAGACGAGGACCTTGGTCTTTTCCCTTGTGTTTAAAGCTTTGGGCTAGT 379
Db 174 CTTATGGGCGCTCGGAGCCAGATCTCGGCTGCCCCCTGTTGAGGCTTTGGGCTAGT 233
QY 380 GGATCACCAGCTGTTGCTGTTCTATGATGATGAGAGTCCGCTGTGGAGCCCGCAACTGCC 439
Db 234 GGACGACCAGCTGTTGTTGTCTCTACAATCAGAGAGTCCGCTGCAGAGCCTCGCGCCC 293
QY 440 ATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGCTGAGTCAAGAGTCTGAAAGG 499
Db 294 CTGGCTCTCGGTAAGGCTCCACCAACTGTGGCTGCAGTAAGCCAGACCTGAAAGG 353
QY 500 GTGGATCACATGTTCACTGTGACTTCTGCACTATTATGAAATATCAACACCACAGCAA 559
Db 354 GTGGATCACATGTTCACTGTGACTTCTGCACTATTATGAAATATCAACACCACAGCAA 413
QY 560 G-----GAGTCCCAACACCTGCAGCTCATCTCTGGGCTGTGA 595
Db 414 GATAACGAGCTGGAGTGTGGCAGAGTCCCAACCTCCAGGTGATCTTGGGCTGTGA 473
QY 596 AATGCAAGAAGACACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCGCAGGACCA 655
Db 474 AGTGAAGCGGACACACAGCACCAGAGGGTCTTGAAGTATGGGTACGATGGCGCAGGACCA 533
QY 656 CTTTGAATTTCTGCCTG 672
Db 534 CTTGAGTTCCACCCCTG 550

RESULT 6
BF080089
LOCUS 230846 MARC 2PTG Sus scrofa cDNA 5', mRNA linear EST 18-OCT-2000
DEFINITION 523 bp
ACCESSION BF080089
VERSION BF080089.1 GI:10873919
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 523)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGAGG
Plate: 100 row: C column: 24
Seq primer: ATTTAGGTGACACTATAG.

```

FEATURES source Location/Qualifiers

1. .523

/organism="Sus scrofa"

/db_xref="taxon:9823"

/clone_lib="MARC 2Pig"

/tissue_type="pooled"

/lab_host="DH10b"

/note="Vector: pCMV SPOR6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from testis, ovary,

endometrium, hypothalamus, pituitary, and placenta."

103 a 175 c 152 g 93 t

BASE COUNT

ORIGIN

Query Match 19.1%; Score 275; DB 10; Length 523;

Best Local Similarity 74.0%; Pred. No. 1.2e-65;

Matches 387; Conservative 0; Mismatches 110; Indels 26; Gaps 2;

QY 119 AGCCTAGGCAATAGCTGTAGGCTGACTCTTGAGGACCATCCCGTTTCCCGCGCCCAAA 178

DB 2 AGCCTTGGCAATGCTCCAGGCTGACTCTTGAGCCACCTCGGTTTCCCGCGCCCA 61

QY 179 AGAAGCGGAGATTAAAGGGGACGTGGCGGCAGAGCTGGGGAATAGGCGCGGAGCCAG 238

DB 62 AGCGCGCGAAAGCACCTGGAAGCGCGGTCCGAGCC--GGGAATAGGCGCCACAGGCCG 119

QY 239 GCCGGCGCTTCTCCTCTGATGCTTTTGACAGCGCGGTCTGCGAGGGCGCTTGCTGCG 298

DB 120 GCCGGCGCTTCTCCTCTGATCCTCTGCGGACCGTGGCCACGCATGATGAGCGCGCG 179

QY 299 TTCACACTCTCTGCACCTACCTTTTCATGGGTGCTCAGAGCAGGACCTTGGTTCCTT 358

DB 180 GCCACACTCCCTGCTCTTCTCTTCTATGGGCGCTCGAGAGCAGATCGGGTGCCTCT 239

QY 359 GTTTGAAGCTTTGGCTACGCGGATGACCAAGCTGTCGTGCTTCTATGATGATGAGATCG 418

DB 240 GTTTGAGCGCTTGGGCTACGTGGAGACCACTGTTTGTCTTCTACATCAGAGAGTCG 299

QY 419 CCGTGTGAGAGCCCGAACTCCATCGGTTTCAGTAGAATTTCAAGCCAGATGTGGCTGCA 478

DB 300 CCGTGCAGAGCTCGCGCCCTGGCTCTCGGGTAAGCGCTTCCAAACCAACTGTGGCTGCA 359

QY 479 GCTGAGTCAGAGCTGTAAGAGGTGGATCAGATCTTCACTGCTTCACTTCTGGACTATTAT 538

DB 360 GCTAAGCAGAGCTGTAAGAGGTGGATCAGATCTTCACTGCTTCACTTCTGGACCATCAT 419

QY 539 GGAAATCACAAACACAGCAAG-----GAGTCCCCACACCT 574

DB 420 GGACAACCAACATACACGAAGTAACGAAGCTGGGAGTGTGGCAGAGTCCCCACACCT 479

QY 575 GCAGTGCATCTGGGCTGTGAATTCGAAGAGACAAACAGTACC 617

DB 480 GCAGGTGATCTGGGCTGTGAAGTGAAGCGGACAAACAGCAC 522

RESULT 7

BE994943 489 bp mRNA linear EST 05-OCT-2000

LOCUS UI-M-CG0p-b1k-d-03-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone

DEFINITION UI-M-CG0p-b1k-d-03-0-UI 3', mRNA sequence.

ACCESSION BE994943

VERSION BE994943.1 GI:10678689

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

TITLE 1 (bases 1 to 489)

ABSTRACT Ronaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

```

QY 694 CAGAAC 700
Db 483 CCAGCC 489

RESULT 8
LOCUS A1850020 457 bp mRNA linear EST 15-JUL-1999
DEFINITION UI-M-BG0-aib-g-10-0-UI.s1 NIH_BMAP_MSC Mus musculus cDNA clone
ACCESSION UI-M-BG0-aib-g-10-0-UI 3', mRNA sequence.
VERSION A1850020
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 457)
JOURNAL Bonaldo,M.F., Lennon,G. and Soares,M.B.
MEDLINE Normalization and subtraction: two approaches to facilitate gene
COMMENT discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: NIH BMAP cDNA clones will be made
available by the means that is soon to be determined. when NIH
determines the means for distribution of the BMAP cDNA clones, this
record will be updated accordingly when that means is determined.
The following repetitive elements were found in this cDNA sequence:
3-30. >(CAG)n$Simple_repeat
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
Location/Qualifiers
1..457
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BG0-aib-g-10-0-UI"
/clone_lib="NIH_BMAP_MSC"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3P-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_MSC library is a non-normalized library
constructed from mouse spinal cord. The tag is a string
of 5 nucleotides present between the Not I site and the
oligo-dT track. The library was constructed as described
by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
1996. Tissue provided by Ms. Annie Novakovich,
Zivic-Miller Laboratories.
TAG_LIB=NIH_BMAP_MSC
TAG_TISSUE=spinal-cord
TAG_SEQ=TCGAA"

BASE COUNT 100 a 129 c 144 g 84 t
ORIGIN

Query Match 17.5%; Score 251.4; DB 9; Length 457;
Best Local Similarity 79.6%; Pred. No. 4.4e-59;
Matches 297; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 561 GAGTCCACACCTCGAGGTCACTCTGGGCTGTGAATGCAAGAGACACAGTACCGAG 620
Db 85 GAGTCCACACCTCGAGGTGGCTCTAGGCTGTGAGGTGCTGAAGACACAGTACCGAG 144

```

```

QY 621 GGCTACTGGAAGTACGGGTATGATGGGAGGACACCTTGAATTCTGCCCTGCACACTG 680
Db 145 GGCTTCTGGAGATATGGTTATGACGGGCAAGATCACCTTGGAAATCTGCCCAAGACACTA 204
QY 681 GATTGGAGAGCAGACAGAACCCAGGGCCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAG 740
Db 205 AACTGGAGCGCAGCGAGCCAGGGCCTGGCCACCAAGTGGAAATGGGAGCAGCACAAG 264
QY 741 AATCGGGCAGGACGACAGACAGGCGCTACCTGGAGAGGGACTGCCCTGCACACTGCAGCAG 800
Db 265 ATCCGTGCCAAACACAGACAGGAGCTACCTGGAGAGGAGTCCGCCCGAGCAGCTGAAACGG 324
QY 801 TTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGCCTCTTGTGGTGAAGTGAACA 860
Db 325 CTCCTGGAGCTGGGAGAGGCGTCTGGACACAGCAAGTGCCTACTTTGGTGAAGTGAACA 384
QY 861 CATCATGTGACCTCTCAGTACCACTTACGGTGTGGGCGCTTGAAGTACTATACCCCGAG 920
Db 385 CGCCACTGGGCGCTCTACGGGAGCTCTCTAAGGTGTGAGGCTCTGGAGCTTCTTCCCGCAG 444
QY 921 AACATCACCATGA 933
Db 445 AACATCATTATGA 457

RESULT 9
LOCUS BE995172 455 bp mRNA linear EST 05-OCT-2000
DEFINITION UI-M-CG0p-bil-h-10-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
ACCESSION BE995172
VERSION BE995172.1 GI:10679153
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 455)
JOURNAL Bonaldo,M.F., Lennon,G. and Soares,M.B.
MEDLINE Normalization and subtraction: two approaches to facilitate gene
COMMENT discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
The following repetitive elements were found in this cDNA sequence:
1-31. >(CAG)n$Simple_repeat
Seq primer: M13 Forward
POLYA=No.

FEATURES
source
Location/Qualifiers
1..455
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-CG0p-bil-h-10-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3P-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NIH_BMAP_Ret4_S2 library is a subtracted library,

```


fax: 402.702.4350
 Email: smith@email.marc.usda.gov
 Single pass sequencing, Bases called and alt_trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -mnscore 18
 and -minmatch 12 options.

Email: Smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
V0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCACGACG
Plate: 71 row: A column: 5
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1. .546

```

/organism="Bos taurus"
/ab_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="vector: pCMV SP0RT6; Site_1: XbaI; Site_2: XhoI;
Library made from pooled tissue from testis, thymus,
semitendinosus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
126 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200
BASE COUNT      a      c      g      t      others
126 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200

```

BASE COUNT	126 a	139 c	180 g	100 t	1 others
ORIGIN					.

Query Match	14.9%	Score 214.6;	DB 10;	Length 546;
Best Local Similarity	85.9%;	Pred. No. 9.2e-49;		
Matches 238;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;
QY	561	GAGTCCCAACCCCTGCAGGTGATCCTGGGTGTGTAATGCAAGAACACCAACAGTACCGAG	620	
Db	139	GAGTCCCAACCCCTGCAGGTGATCCTGGGTGTGAGTGTGCAGAGGACACACACCCAGA	198	
QY	621	GGCTACTGGAAGTACGGGTATGATGGCAGGACACCACTTGAATTCGCCCTGCACACATG	680	
Db	199	GGGTTCTGGAAGTACGGGTACGATGGCAGGACCACTTGAATTCGGGCTGAGACATG	258	
QY	681	GATTGGAGAGCAGACAGAACCCAGAGGCGCTGGCCCAACCAAGCTGCAGTGGGAAGGACCAAG	740	
Db	259	GATTGGAGAGCAGACAGCAACCCAGAGGCCCAAGTCAACCAAGCTGGAGTGGGAGTGCACAG	318	
QY	741	ATTTCGGGCCAGCGACAAACAGGGCGCTACTGGAGAGGAGCTGCCCTGCACAGTTCGACAG	800	
Db	319	ATTTCGGGCCAAGCAAAACAGAGGCGCTAAGTGGATCGGATTGCCCGAGGAGCTGGTGCAT	378	

Db 379 TTGCTGGAGCTGGGGAGAGGGCCTCTGGAGCAGCAAG 415

```

RESULT 13
BB851691
LOCUS BB851691 481 bp mRNA linear EST 26-NOV-2001
DEFINITION BB851691 RIKEN full-length enriched, B16 F10Y cells Mus musculus
cDNA clone G370002P09 5', mRNA sequence.
ACCESSION BB851691
VERSION BB851691
KEYWORDS BB851691.1 GI:17093145
SOURCE EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 481)
AUTHORS Akimura,T., Arakawa,K., Carninci,P., Furuno,M., Hanagaki,T., Ishii
Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.,
Saito,R., Sakai,K., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.,
Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES
source Location/Qualifiers
1. .481
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="G370002P09"
/clone_lib="RIKEN full-length enriched, B16 F10Y cells"
/cell_type="B16 F10Y cells"
95 a 134 c 140 g 112 t
BASE COUNT
ORIGIN

Query Match 14.6%; Score 209.6; DB 9; Length 481;
Best Local Similarity 68.5%; Pred. No. 2.1e-47;
Matches 307; Conservative 0; Mismatches 134; Indels 7; Gaps 1;

Qy 114 TCCTGACCTAGGCAATAGCTGTAGGCTGACTTCTCGAGCCATCCCGTTTCCCGCCCC 173
11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 9 TCTGGACCTCAGCAATGGCTACAGGCTGACTTCTTGGATCCTCCACGTTTCCAGATCCT 68

```

```

QY 174 CCAAAAGAGCGGAGATTTTAACGGGACCTGCGGCCAGAGCTGGGGAATGGGCCCGCA 233
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 69 AGTGAAGACCGGTGGACCCAGC-----TGAGGACATGAGCCTATCAGCTGGGCTCCCT 121
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 234 GCCAGCGCGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTCTGAGGGGCGCTTG 293
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 122 GTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 294 CTGCGTTACACTCTCTGCACTACCTCTTTCATGGGTGCTCAGAGCAGACCTTGGTCTT 353
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 182 CCCTGTTACATCTCTTAAGATACCTCTTTCATGGGTGCTCAGAGCAGACCTCGGCTG 241
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 354 TCCTGTTTGAAGCTTTGGCTACGTGGATGACCACTGCTGCTGCTGCTGCTGCTGCTGCTG 413
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 242 CCTTTGTTGAGCTAGGGGCTATGTTGATGACCACTCTTTGTGCTTCCATCATCATGAG 301
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 414 AGTCGCGGTGTGGAGCCCCCACTCCATGGGTTCAGTAGAATTTTCAAGCCAGATGTGG 473
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 302 AGTCGCGGTGTGGAGCCCCCGTGGATCTTGGAGCAAACTCAAGCCAGCTGTGG 361
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 474 CTGAGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCTACTGTTGACTTCTGGACT 533
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 362 CTGCATCTGAGTCAGAGCTTGAAGGGTGGGACTACATGTTTCATAGTAGACTTCTGGACC 421
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
QY 534 ATTATGGAATCACACACCAAGG 561
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 422 ATCATGGGCAACTATAACCCAGTAAGG 449
1111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
RESULT 14
BI452668
LOCUS BI452668 831 bp mRNA linear EST 21-AUG-2001
DEFINITION 603169877F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5249395 5',
mRNA sequence.
ACCESSION BI452668
VERSION BI452668.1 GI:15243324
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 831)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11629 row: j column: 20
High quality sequence stop: 818.
FEATURES
source Location/Qualifiers
1. .831
/organism="Mus musculus"
/strain="C57/B6"
/db_xref="taxon:10090"
/clone="IMAGE:5249395"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: Sali;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"
BASE COUNT 207 a 220 c 189 g 215 t
ORIGIN

```

THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222

Search completed: June 19, 2002, 07:21:51
Job time: 2597 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: June 19, 2002, 08:49:52 ; Search time 352.95 Seconds
(without alignments)
7004.836 Million cell updates/sec

Title: US-09-497-957-12
Perfect score: 1440
Sequence: 1 GGGGACACTGGATCACCTAG.....TCCTCAAAAGATTTCCTCCA 1440

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_032802.*			
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*	1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*	1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*	1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*	2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*	2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*	2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*	3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*	3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*	3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*	4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*	4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*	4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*	5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*	5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*	5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*	6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*	6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*	6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*	7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*	7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*	7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*	8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*	8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*	8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*	9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*	9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*	9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*	10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*	10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*	10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*	11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*	11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*	11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*	12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*	12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*	12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*	13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*	13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*	13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*	14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*	14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*	14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*	15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*	15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*	15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*	16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*	16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*	16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*	17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*	17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*	17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*	18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*	18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*	18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*	19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*	19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*	19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*	20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*	20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*	20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*	21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*	21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*	21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*	22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*	22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*	22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*	23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*	23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*	23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*	24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*	24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*	24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES							
Result No.	Score	Match	Length	ID	Description		
1	1440	100.0	1440	22	AAC68432		Human hereditary h
2	1438.4	99.9	1440	22	AAC68430		Human hereditary h
3	1436.4	99.9	1440	22	AAC68431		Human hereditary h
4	1436.8	99.8	1440	18	AAT96691		Hereditary haemoch
5	1436.8	99.8	1440	22	AAC68429		Human hereditary h
6	1436.8	99.8	2727	19	AAV23525		Haemochromatosis g
7	1215.8	84.4	2506	21	AA96769		CDNA sequence enco
8	557.6	38.7	596	22	AA163897		Human polynucleoti
9	321	22.3	10825	22	AAC68427		Human hereditary h

10	321	22.3	10825	22	AAC68428		Human hereditary h
11	319.4	22.2	10825	18	AAT96690		Hereditary haemoch
12	319.4	22.2	10825	22	AAC68425		Human hereditary h
13	319.4	22.2	10825	22	AAC68426		Human hereditary h
14	319.4	22.2	12146	21	AA96794		Genomic DNA of a h
15	319.4	22.2	23736	19	AAV57903		Hereditary haemoch
16	280	19.4	517	22	AAC68441		Human hereditary h
17	278.4	19.3	359	20	AA160555		Human hereditary h
18	278.4	19.3	517	22	AAC68440		Human hereditary h
19	278.4	19.3	5749	22	AA136747		Human hereditary h
20	191.8	13.3	8622	24	ABL34142		Human immune syste
21	178.2	12.4	8622	24	ABL34143		Human immune syste
22	173.6	12.1	1112	21	AAA48668		CDNA encoding chic
23	172	11.9	1230	21	AAA48673		CDNA encoding chic
24	168.8	11.7	1195	21	AAA48671		CDNA encoding chic
25	166.8	11.6	1197	21	AAA48672		CDNA encoding chic
26	166.8	11.6	1262	21	AAA48691		Consensus DNA of c
27	164	11.4	1173	21	AAC78071		Human cancer assoc
28	164	11.4	1173	22	AAF72767		Human prostate can
29	162	11.2	1145	21	AAA48667		CDNA encoding chic
30	160.4	11.1	1230	21	AAA48669		CDNA encoding chic
31	160.4	11.1	1284	9	AA80603		Probe F10 of Major
32	158.8	11.0	1230	21	AAA48665		CDNA encoding chic
33	158.8	11.0	1230	21	AAA48670		CDNA encoding chic
34	157.2	10.9	1230	21	AAA48666		CDNA encoding chic
35	155.6	10.8	1230	21	AAA48664		CDNA encoding chic
36	143.6	10.0	1554	22	AA193004		Human polynucleoti
37	143.2	9.9	1101	12	AA012116		HLA-C exon Cb-1
38	142	9.9	1567	22	AAH98676		Human EST-derived
39	141.8	9.8	1073	22	AAH42223		Nucleotide sequenc
40	141.2	9.8	1098	22	AA07697		Human CDNA encodin
41	140	9.7	1101	12	AA012117		HLA-C exon Cb-2.
42	139.4	9.7	2034	23	AA590913		DNA encoding novel
43	139.4	9.7	2037	23	AA590740		DNA encoding novel
44	138.4	9.6	4965	16	AA075973		pHLA-B7/beta-2 mic
45	137.8	9.6	1089	22	AAH45555		Human cancer cell

ALIGNMENTS

RESULT	1
AAC68432	
ID	AAC68432 standard; DNA; 1440 BP.
XX	
XX	AAC68432;
XX	
XX	21-FEB-2001 (first entry)
DE	
XX	Human hereditary hemochromatosis 24d1/2 mutation cDNA.
XX	
KW	HH; hereditary hemochromatosis; chelation agent;
KW	T-cell differentiation factor; iron overload; ss.
XX	
OS	Homo sapiens.
XX	
XX	US6140305-A.
XX	
PD	31-OCT-2000.
XX	
PF	04-APR-1997; 97US-0834497.
XX	
PR	04-APR-1996; 96US-0630912.
PR	16-APR-1996; 96US-0632673.
PR	23-MAY-1996; 96US-0652265.
XX	
PA	(BIRA) BIO-RAD LAB INC.
XX	
PI	Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI	Feder JN;
XX	
DR	WPI; 2001-006341/01.
XX	

PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS
PS
XX
XX
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
XX
SQ Sequence 1440 BP; 348 A; 354 C; 407 G; 331 T; 0 other;

Query Match 100.0%; Score 1440; DB 22; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
Db 1 9999aacactggatacactagtgtttcacaaagcaggtacottctgtgttaggaagagaga 60
QY 61 ACTAAAGTTCTGAAGACCTGTGCTTTTCACAGAAAGTTTACTGGGCATCTCCTGAG 120
Db 61 actaaagtctgaagacacctgtgcttttcaccaggaaagttttactgggcatctctctgag 120
QY 121 CCTAGGCAATAGCTGTAGGCTGACCTCTGAGCCATCCCGCTTCCCGCCCCCAAG 180
Db 121 cctaggcaatagctgtagggtgactctgtgagccatcccgcttccccgcgcccccaag 180
QY 181 AAGCGAGATTTAAACGGGACCTGCGGCCAGAGCTGGGAAATGGCGCCGAGCCAGGC 240
Db 181 aagcgagatttaacgggagactgctggccagagctgggaaatggccgcagccaggc 240
QY 241 CGCGCTTCTCCCTCCATGCTTTTTCAGACCGCGCTCTCAGGGCGCTTGTGCGT 300
Db 241 cgcgcttctccctccatgcttttgcagaccgctgctgagggcgcttgcgt 300
QY 301 CACACTCTCGCACTACCTCTTCATCGGTGCTCAGAGCAGACCTTGGTCTTCTTCT 360
Db 301 cacactctctgactacctctctatgggtgcctcagagcagaccttggctctctctgt 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACAGCTGTTCTGTTTATGATGATGAGCTCGCC 420
Db 361 ttgaagctttgggctacgtggatgaccagctgttctgttctatgatgatgagctgcgc 420
QY 421 GTGTGAGCCCCGAACTCCATGCGTTTCCAGTAGAATTTCAAGCCAGATGTGCTGCAGC 480
Db 421 gtgtgagccccgaaactccatgggtttccagtagaatttcaagccagatgtggtgcagc 480
QY 481 TGAGTCAGACTGGAAGGTGGGATCACATGTTCACTGTGACTTCTGACTATTATGG 540
Db 481 tgaagtcagactgcgaaggtgggaaccatgttccactgttgaactcttggactattatgg 540
QY 541 AAAATCACACACACAGAGAGTCCACACCTGTCAGAGTTCATCCCTGGGCTGTGAATGC 600
Db 541 aaaatcacacacacagagagtcacacacctgcaggtcatcctggctgtgaaatgc 600
QY 601 AAGAAGACAACAGTACGAGGGCTACTGGAGTACGGGTATGATGGCAGGACCCTTTC 660
Db 601 aagaagacaacagtcacgagggctacttgaagtcaggggtatgatggcagggaccacttg 660
QY 661 AATGCTGCCCTGACACTGATTTGGAGAGCAGCAACCCAGGGCTGCGCCACCAAGC 720
Db 661 aatctgccttgacacactggaattggagagcagcaacccagggccttggcccccaagc 720
QY 721 TGGAGTGGGAAAGCACAAGATTTCGGGCCAGGACAGGGGCTTACCTGGAGAGGACT 780
Db 721 tggagtgggaaagcacaagatttcggccaggcagaaagggcctacctggagagggact 780
QY 781 GCCCTGCACAGCTGCAGCAGTTCCTGGAGCTGGGGAGAGGTGTTTTGGACCAACAAGTGC 840

Db 781 gccctgcacagctgcagcagttctgtggagctggggagaggtgttttggaccaacaagtgc 840
QY 841 CTCTCTTTGGTGAAGGTGACACATCATGTGACCTTTCAGTGACCCTCTACGCTGTCGGG 900
Db 841 ctctcttttgggaaggtgacacatacgtgacctctctcagtgaccactctacggtgctcgg 900
QY 901 CCTTGAAGTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
Db 901 ccttgaagactactaccctcccaagaacataccatgaagtgtcgtgaaggataagcagccaatgg 960
QY 961 ATGCCAAGAGGTTCGAACCTTAAAGACCTATTGCCCCAATGGGGATGGACCTACACAGGCT 1020
Db 961 atgccaaaggagttcgaacacctaaagacgtattgtcccaatggggatgggacacacagggct 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGCAGAGATATACGTACAGGTGGAGCACC 1080
Db 1021 ggaataaccttggctgtacctccctggggagagcagagatatatacgtaccaggtggagcacc 1080
QY 1081 CAGGCCCTGGATCAGCCCTCATTTGTATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db 1081 caggccctggatcagccctcatctgtgatctgggagccctcacctgtcggcaccctagtc 1140
QY 1141 TTGGAGTCATCAGTGAATTCGCTGTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA 1200
Db 1141 ttggagtcatcagtggaattgctgtgttctgtcgtcctctgttcttctggaatttgttca 1200
QY 1201 TAATATTAAAGGAAGGAGCAGGGTTCAAGAGAGGAGCCATGGGSCACTACGTCTTACGTGAAC 1260
Db 1201 taatattaaggaagagcaggggttcaagagagccatggggcactacgtctctagctgaac 1260
QY 1261 GTGAGTCACAGCCGCTCGACACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 gtgagtcacagcgcgcctgcagactcactgtgggaaggagacacaaactagagactcaag 1320
QY 1321 AGGGAGTCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 agggagtcattatgagctcttcattgttcttcagagagagcttgaacctaaacatagaaat 1380
QY 1381 TGCCTGACGAAGTCTCTGATTTAGCCCTTCCTCTTTCATTTCCCTCAAAAAGATTTCCCA 1440
Db 1381 tgcctgacgaagctccttgatttagccttctctgttcttctcctcaaaaagatttcccca 1440

RESULT 2
AAC68430
ID AAC68430 standard; DNA; 1440 BP.
XX
AC AAC68430;
XX
DT 21-FEB-2001 (first entry)
XX Human hereditary hemochromatosis 24d1 mutation cdNA.
XX
XX HH: hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ss.
XX Homo sapiens.
XX
XX US6140305-A.
XX
XX 31-OCT-2000.
XX
XX 04-APR-1997; 97US-0834497.
XX
XX 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
XX (BIRA) BIO-RAD LAB INC.
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;

XX WPI: 2001-006341/01.

XX New hereditary hemochromatosis gene products or polypeptides, useful

PT for treating hereditary hemochromatosis in a patient, and as a metal

PT chelation agent alleviating iron overload -

XX

XX Disclosure; Fig 4; 108pp; English.

XX The present invention relates to hereditary hemochromatosis gene

CC products. These proteins may be used to treat a patient diagnosed as

CC having human hemochromatosis disease. It is also useful as a metal

CC chelation agent or as a T-cell differentiation factor, and for

CC alleviating iron overload. They may also be used in protein replacement

CC therapy for individuals having a defective human hemochromatosis gene.

XX

SQ Sequence 1440 BP; 348 A; 355 C; 406 G; 331 T; 0 other;

Query Match 99.9%; Score 1438.4; DB 22; Length 1440;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATACCTAGTGTTCACAGCAGGTACCTTCTGCTAGGAGAGAGA 60

DB 1 9999acactggatcacctagtggttcacagcaggtaccttctgtctagagagaga 60

QY 61 ACTAAGTCTCAAGACACTGTGCTTTTCCACAGGAGTCTTACTGGGATCTCTCTAG 120

DB 61 actaagttctgaagacactgtgcttttccacaggaagtttactgggactctctctgag 120

QY 121 CCTAGGCAATAGCTAGGGTGACTTCTGGAGCCATCCCGCTTCCCGCGCCCAAAAG 180

DB 121 cctaggcaatagctaggggtgaactcttgagaccatcccgcttcccgcccccaaaag 180

QY 181 AAGCGGAGTTTAAGCGGAGCTGGCGCCAGAGCTGGGGAATGGCCCGCGACCGAGCC 240

DB 181 aagcggagatttaacg99agc9gc99ccagagctg999aaatg99cc9c9agc9ag9c 240

QY 241 CGGCGCTTCTCTCTGTATGCTTTTTCAGACCGCGGCTCTGTCAGGGCGCTTGTCGGTT 300

DB 241 cggcgcttctctctgtatgcttttgcagac9cg9gctctgcagggcgcttgc9gctt 300

QY 301 CACACTCTCTGCACATCTCTTCATGGTGTCTCAGAGCAGGACCTTGGTCTTCTCTGT 360

DB 301 cacactctctgcactctctctcaatg99gtcctcagagcaggaacttggctcttctctgt 360

QY 361 TTGAAGCTTTGGGCTACGCTGGATGACAGCTGTTCGTGTCTATGATGATGAGAGTCGCC 420

DB 361 ttgaagctttgggctacgtggatgaccagctgttctgttctatgatcatgagagtcgcc 420

QY 421 GTGTGAGCGCCGAACCTCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

DB 421 g9t9g9g9cc9gaactcca99gttccagtagaattccaagccagatg9t9g9c9gagc 480

QY 481 TGAGTCAGAGTCTGAAGGGTGGATACATGTTCACTGTTGACTTCTGGACTATTATGG 540

DB 481 tgagtcagagctctgaaggg9tggatcacatgttcaactgttgacttcttgactattatg 540

QY 541 ARAATCACACACAGCAGGAGTCCACACCTGCAGTCTATCTCGGCTGTGAATGC 600

DB 541 aaaaatcacacacagcagga99gttcccaacaccc9caggtctatcctg99c9t9gaatgc 600

QY 601 AAGAGACAACAGTACCGAGGCTTACTGGAAGTACGGGTATGATGGGACGACCACTTG 660

DB 601 aagaagacaacagtaccagggctactggaagtacgggtatgatgggacgacaccttg 660

QY 661 AATTCGCGCTGCACACTGGATTTGGAGCAGCAGACCCAGGCGCTGGCCACCAAGC 720

DB 661 aattctg9cctgacacactg9at9gagcagcaacccag9g9cctg9cc9cccaagc 720

QY 721 TGGAGTGGGAAAGGACAGATTCGGGCCAGGCAGAGACGGGCTTACTGTGAGAGGACT 780

DB 721 tggagtgggaaaggcacagaagattcg9ccagggcagaacacagg9cctacctggagagggact 780

QY 781 GCCCTGCACACCTGCAGAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840

DB 781 gccctgcacagctgcagcagctgctggagctg999gag9g9ctttt99acccaaga9tgc 840

QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACACTCTACGGTGTGGG 900

DB 841 ctcctttgtgaaggtgacacatcatgtgacctcttcagtgaccactctacgggtgctcgg 900

QY 901 CCTTCAACTACTACCCAGAACATACCATGAAGTGGCTGAAGGATAAGCACCAATGS 960

DB 901 ccttgaaactactacccacagaaacatccca9aag99c9ga99gata9a9cagcaat9g 960

QY 961 ATGCCAAGAGTTCGAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACCAAGGCT 1020

DB 961 atgccaagagttcgaacctaaagacgtattgcccaatgggattgggacctaaccaggct 1020

QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATAGTACCAGGTGGACACC 1080

DB 1021 ggataaccttggctgtacctccctg99gaagagcagagatatagctaccaggtg9agcacc 1080

QY 1081 CAGGCTCGATCAGCCCTCATTTGTGATCTGGAGCCCTCACCGTCTGGCACCTTAGTCA 1140

DB 1081 caggcctggatcagccctcatgtgatctgggagccctcaccgctc9gcacccctagtca 1140

QY 1141 TTGGAGTCATCAGTGGAAATGCTGTTTTCGTCATCTTGTTCATTGGAAATTTGTTCA 1200

DB 1141 ttggagtcatacagtg99aat9c9t9ctt9ct9c9c9ct9c9c9c9c9c9c9c9c9c9c9 1200

QY 1201 TAATATTAAGGAAGAGCAGGGTTCAAGAGAGGCCATGGGGCAGCTACGTTTACGTGAAC 1260

DB 1201 taatattaaggaagagggcagggttcaagagagggccatggggcactacgtctctagctgaac 1260

QY 1261 GTGAGTGACACGAGCCCTGCAGACTCTGTGGGAAGAGACAAACTAGAGACTCAAAG 1320

DB 1261 gtgagtgcacagc9 1320

QY 1321 AGGAGTGCATTTATGAGCTCTTTCATGTTTCAGGAGAGTTCGAACCTAAACATAGAAAT 1380

DB 1321 agggagtgcat9c9 1380

QY 1381 TGCCCTGAGCACTCCCTGATTTTACGCTTCTCTGTTTTCATTTCTCAAAAAGATTTCCCCA 1440

DB 1381 tgccctgacgaactcctgatttttagc9c9c9c9c9c9c9c9c9c9c9c9c9c9c9c9c9c9 1440

RESULT 3

AAC68431

ID AAC68431 standard; DNA; 1440 BP.

XX AAC68431;

XX AC AAC68431;

XX DT 21-FEB-2001 (first entry)

XX DE Human hereditary hemochromatosis 24d2 mutation cDNA.

XX HH; hereditary hemochromatosis; chelation agent;

XX T-cell differentiation factor; iron overload; ss.

XX OS Homo sapiens.

XX PN US6140305-A.

XX PD 31-OCT-2000.

XX PF 04-APR-1997; 97US-0834497.

XX PR 04-APR-1996; 96US-0630912.

XX PR 16-APR-1996; 96US-0632673.

XX PR 23-MAY-1996; 96US-0652265.

XX PA (BIRA) BIO-RAD LAB INC.

XX Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 4; 108pp; English.
XX The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
XX Sequence 1440 BP; 347 A; 354 C; 408 G; 331 T; 0 other;
SQ

Query Match 99.9%; Score 1438.4; DB 22; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGTACCTTCTGCTAGGAGAGAGAGA 60
DB 1 ggggacactggatcacctagtgtttcacagcaggtacacctctgtgtaggagagaga 60
QY 61 ACTAAAGTTCTGAAGACCTGTGCTTTTACAGGAAAGTTTACTGGGCATCTCCTGAG 120
DB 61 actaaagtctgaagacctgtgtctttcacaggaagtattactgggcatctcctgag 120
QY 121 CCTAGGCAATAGCTGAGGTGATCTTCTGAGCCATCCCGCTTCCCGCCCGCCCAAGAG 180
DB 121 cctaggcaatagctgaggtgactcttgagccatcccgcttcccgcccgcccaaaag 180
QY 181 AAGCGGAGATTTAACGGGACGTCGGCCAGAGCTCGGAAATGGCCCGCAGCCAGCAGC 240
DB 181 aagcggagatttaacgggacgtgcggccagagctcg99cgaagt99gaaatc99cccgagccag9c 240
QY 241 CGGCGCTTCTCCCTGATGCTTTTTCAGACCCGCTGCTGAGGGCGCTTCTGCTGCTT 300
DB 241 cggcgcttctccctgactgtctttcagaccgcggtcctgcagggcgcttgctgcgtt 300
QY 301 CACACTCTCTGCACTACTCTTCTATGGTGCCCTCAGACAGGACCTGGTCTTTCCTTGT 360
DB 301 cacaactctctgcactactctctcattggtgcctcagagcaggacccttggcttctccttgt 360
QY 361 TTGAAGCTTTGGGCTAGCTGATGACAGCTGTTCTGTTCTATGATGATGAGATCGCC 420
DB 361 ttgaagctttgggctagctgagtgaccagctgttcgtttctatgatgatgagagtcgcc 420
QY 421 GTGTGGAGCCCCAACCTCCATGGTTTCCAGTAGAATTTCAAGCCAGATGTGCTGCAGC 480
DB 421 gtgtggagccccaaactccatgggtttccagtagaatttcaagccagatgtggtcgcagc 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTGTTGACTTATTATGG 540
DB 481 tgagtcagagctgaaggggtg9gatcacatgttcactgttcacttctggacttattatg 540
QY 541 AAAATCAACACACAGAGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAATGC 600
DB 541 aaaatcaaacacagcaagaggtccacacccctgcaggtcactcctgggtgctgaaatgc 600
QY 601 AAGAGACAACAGTACCGAGGGCTTCTGGAAGTACGGGTATGATGGCAGGACCCCTTG 660
DB 601 aagaagacaacagtaccagggcttactggaagtacg9gtgatg9ggcagaccacccctg 660
QY 661 AATTCTGCCCTGACACTGATTTGGAGAGCAGCAGACCAACCCAGGCGCTGGCCCAAGC 720
DB 661 aattctgccctgacacactgattggagagcagcagaaccacggcctgcccaccaaagc 720

QY 721 TGGAGTGGAAAGGCACAAAGATTTCGGGCCAGGACAGAGGCGCTACCTGGAGAGGAGCT 780
DB 721 tggagtggaaaggcacaaagatttcggggccaggcagaaagggcctacctggagaggact 780
QY 781 GCCTTGACAGCTGACAGCTGCTGAGCTGGAGTGGTGTGTTGGACCAACAGATGC 840
DB 781 gccttgacagctgacagctgctgagctgggagaggtgttttggaccacaagatgc 840
QY 841 CTCTTTGGTCAAGGTGACACATCATGTGACCTTTCAGTGACCACTTACGGTGTGCGG 900
DB 841 ctctttgggcaaggtgacacatcatgtgaccttctcagtgaccactctacggtgtcgg 900
QY 901 CTTTGAATCTACTCCCCCAGACATCACCATGAAGTGGCTGAAGGATAGACACCAATGG 960
DB 901 ctttgaactactccccagacatcacccatgaagtgtgctgaaggataagcagccaatgg 960
QY 961 ATGCCAAGGAGTTCGAACCTTAAAGACGTATTGCCCATGGGATGGACCTTACCAAGGCT 1020
DB 961 atgccaaaggagttcgaacctaaagacgtattgcccattgggagctcctacacacagggct 1020
QY 1021 GGATAACCTTGGCTGTACCCCTCGGGAAGCAGAGATATAGCTACAGGTGGAGCAC 1080
DB 1021 ggaataaccttggctgtacctccctgggaaagacagagatatagctgcccaggtggagcacc 1080
QY 1081 CAGGCGCTGATPCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
DB 1081 caggcctgatatcagccctcatctgtgctggagccctcaccgtctg9ccaccctagtc 1140
QY 1141 TTGGAGTCATCAGTGGATTCGTTTCTGTCGTCATCTTGTTCATTTGGAAATTTGTTC 1200
DB 1141 ttggagtcatacagtggaattcgtcttttctgctcatcttctgttcattggaattttgtca 1200
QY 1201 TAAATATTAAAGAAAGCAGGCGGTTTCAAGAGAGGACCATGGGGACACTAGCTCTTACTGAAC 1260
DB 1201 taatattaaagaagaggcaggggttcaagagagagccatgggagcactagcttagtgaa 1260
QY 1261 GTGAGTCACAGCAGCCTGACACTCAGTCTGGGAGGAGACAAACTAGAGACTCAAG 1320
DB 1261 gtgagtcacagcagcctgcagactcactgtggagagagacaaactagagactcacaag 1320
QY 1321 AGGAGTGCATTTATGAGCTCTTTCATGTTTCAGGAGAGAGTTGAACCTTAAACATAGAAAT 1380
DB 1321 agggagtgcatattatgactcttcattgttcaggagagagttgaacctaaacatagaaat 1380
QY 1381 TGCCCTGACGAACTCCTTGATTTAGCCCTTCTGTTTCTTCTCAAAAACATTTCCCA 1440
DB 1381 tgcctgacgaaactccttgattttagccttctctctgttcttctcccaaaaagattccccc 1440

RESULT 4
AAT96691
ID AAT96691 standard; cDNA; 1440 BP.
XX
AC AAT96691;
XX
DT 14-APR-1998 (first entry)
XX
DE Hereditary haemochromatosis gene cDNA clone.
XX
KW Hereditary haemochromatosis; metal toxicity; diagnosis;
KW gene therapy; prenatal screening; human; ss.
XX
OS Homo sapiens.
XX
FT Key
FT CDS
FT mutation
FT /tag= a
FT /tag= g
FT /note= "C to G substitution (24d2 mutation)
FT results in His to Asp substitution."
FT variation 414

QY	1141	TTGAGTCATCACTGGAATGCTGTTTGTGCTCATCTTGTTCATTTGGAATTTGTTC	1200
Db	1141		1200
QY	1201	TAATATTAGGAAGAGCGAGGTTTCAAGAGGAGCCATGGGCACTACCTCTTAGCTGAAC	1260
Db	1201		1260
QY	1261	GTGAGTGACACGACCTGSCAGACTCACTGTGGGAAGGAGACAAACCTAGAGACTCAAAAG	1320
Db	1261		1320
QY	1321	AGGAGTGCAATTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT	1380
Db	1321		1380
QY	1381	TGCTCAGCAACTCCCTTGATTTTAGCTTCTCTGTTCATTTTCTCAAAAAGATTTCCCA	1440
Db	1381		1440
RESULT	5		
ID	AAC68429	standard; DNA; 1440 BP.	
XX	AAC68429;		
DT	21-FEB-2001	(first entry)	
XX	Human hereditary hemochromatosis cDNA.		
DE	HH; hereditary hemochromatosis; chelation agent;		
KW	T-cell differentiation factor; iron overload; ss.		
XX	Homo sapiens.		
XX	US6140305-A.		
XX	31-OCT-2000.		
XX	04-APR-1997;	97US-0834497.	
XX	04-APR-1996;	96US-0630912.	
PR	16-APR-1996;	96US-0632673.	
PR	23-MAY-1996;	96US-0652265.	
XX	(BIRA) BIO-RAD LAB INC.		
PA	Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;		
XX	Feder JN;		
XX	WPI; 2001-006341/01.		
XX	New hereditary hemochromatosis gene products or polypeptides, useful		
PT	for treating hereditary hemochromatosis in a patient, and as a metal		
PT	chelation agent alleviating iron overload -		
XX	Disclosure; Fig 4; 108pp; English.		
XX	The present invention relates to hereditary hemochromatosis gene		
CC	products. These proteins may be used to treat a patient diagnosed as		
CC	having human hemochromatosis disease. It is also useful as a metal		
CC	chelation agent or as a T-cell differentiation factor, and for		
CC	alleviating iron overload. They may also be used in protein replacement		
CC	therapy for individuals having a defective human hemochromatosis gene.		
XX	Sequence 1440 BP; 347 A; 355 C; 407 G; 331 T; 0 other;		
XX			
Query Match	99.8%;	Score 1436.8;	DB 22; Length 1440;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 1438;	Conservative	0; Mismatches	2; Indels
		0; Gaps	0;

QY	1	GGGACACTGGATCACCCTAGTGTTCACAAAGCAGGTACCTTCTGCTGTAGGAGAGAGA	60
Db	1		60
QY	61	ACTAAAGTTCTGAAAGACCTGTTGCTTTTCCACGAGGAATTTTACTGGGGCATCTCTGAG	120
Db	61		120
QY	121	CCTAGGCAATAGCTGTAGGTGACTTCTGAGGACATCCCGTTTCCCGCCGCCCAAAAG	180
Db	121		180
QY	181	AAGCGGAGATTAAACGGGACGTGCGGCCAGAGCTGGGAAATGGGCCCGCCGACCCAGGC	240
Db	181		240
QY	241	CGGCGCTTCCTCCTGATGCTTTTGGAGACCGCGGTCTGTCAGGGGCGCTTGTGCTTGT	300
Db	241		300
QY	301	CACACTCTGCTGACTACCTCTTCATGGTGCCTCAGAGCAGGACCTTGTCTTCTTGT	360
Db	301		360
QY	361	TTGAAGCTTTGGGCTACGTGGATGACGAGCTGTGCTTCTATGATGATGAGAGTGGCC	420
Db	361		420
QY	421	GTGTGGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGGGTGCGAGC	480
Db	421		480
QY	481	TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTGTTGCTTCTGACTATTATGG	540
Db	481		540
QY	541	AAATFCAACACACAGCAAGGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAATGC	600
Db	541		600
QY	601	AAGAGACAACAGTACCAGGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG	660
Db	601		660
QY	661	AATCTGCCCTGCACACTGATGGAGTGGAGAGCAGCAACCCAGGCGCTGCGCCACCAAGC	720
Db	661		720
QY	721	TGGAGTGGAAAGGCACAAAGATTGCGGCCAGGCAGAACAGGGCCTACTGTGAGAGGACT	780
Db	721		780
QY	781	GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC	840
Db	781		840
QY	841	CTCCTTTGGTGAAGTGCACATCATGTGACTCTTCACTGACCACCTACTCGGTGTGGG	900
Db	841		900
QY	901	CCTTGAACCTACTACCCCGAGAACATCACATGAAGTGGCTGAAGGATGAAGCAGCAATGG	960
Db	901		960
QY	961	ATGCCAAGAGTTCGAACCTAAAGACGTATTGCCCCAATTGGGGATGGGACCTACCAAGGCT	1020
Db	961		1020
QY	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATACGTACCAGGTGGAGCACC	1080
Db	1021		1080

||||| 361 atcctggctgtgaaatgcaagaacacacagataccgagggctactgaaatcagggat 420
QY 642 GATGGCAGACACACCTTGAAATTCGCCCTGACACACTGGATTGGAGAGCAGCAGACCC 701
Db 421 gatggcagaccacctgaaattgcctgcacactggaattggagagcagagaacc 480
QY 702 AGGCGCTGCCACCAAGCTGGAGTGGGAAGGCACAGATTCGGCCAGCAGACAGG 761
Db 481 agggcctggcccaccagctggatgggaagggcacaagattcgggccaggcagaacagg 540
QY 762 GCCTACTGTGAGAGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGT 821
Db 541 gccactctagagagactgcctgcacagctgcagcagtgctgagactgggagaggt 600
QY 822 GTTTTGGACCAACAGTGCCTGCTTGTGTAAGTTCGACATCATGTGACCTCTTCAGTG 881
Db 601 gtttggaccaacaadgtcctcttggtagaaggtgacacatcatgtgacctcttcagt 660
QY 882 ACCACTCTACGGTGTGCGGCTTGAACCTACTACCCGCCAGAACATCACCATGAAGTGGCTG 941
Db 661 accactctagctgcggcctgaactactacccccagaaacatcacatgaagtgcgtg 720
QY 942 AAGGATAAGCAGCCCAATGATGCCAAGAGTTCGAACCTAAAGACGTATTGCCCAATGGG 1001
Db 721 aaggataagcagccaatgatgccaaaggagttcgaaacctaaagacgtattgcccaatggg 780
QY 1002 GATGGGACCTPACCAGGGCTGGATAACCTTTGGCTGTACCCCTGGGGAAGACAGAGATAT 1061
Db 781 gatgggaacctaccagggctggataaaccttggctgttaccacctggggagagacagagat 840
QY 1062 ACGTACAGGTGAGGACCCAGCCCTGATCAGCCCTCATTTGTGATCTGGGAGCCCTCA 1121
Db 841 acgtgccaggtggagcaccaggcctggatcagccctcattgtgatctgggagccctca 900
QY 1122 CCGTCTGGACCTAGTCAATGTGAGTCATCAGTGAATTCGTTTGTGCTCATCTTG 1181
Db 901 ccgctcggacccttagctccatcggagtcacagtcgaattgctgttttgcgtcatcttg 960
QY 1182 TTCATTGGAATTTGTTCATAATTAAGAGAGAGCGGTTCAAGAGAGCCCATGGG 1241
Db 961 ttcattggaattttgtcataatataaggaagagcaggggttcaagaggagccatgggg 1020
QY 1242 CACTACTCTTAGCTGAACCTGAGTCACACGGCCTGCAGACTCAGTGTGGGAAGGAGA 1301
Db 1021 cactacgctctagctgaacgtgagtgacacgcagcctgcagactcactgtgggaaggaga 1080
QY 1302 CAAACTAGAGACTCAAAGGGAGTGCAATTTATGAGCTCTTTTCATGTTTCAGGAGAGT 1361
Db 1081 caaaactagagactcaagaggagtgcatattatgagctcttcattgttccaggagagagt 1140
QY 1362 TGAACCTAAACATAGAAATTCCTGACGAACTCTTGATTTAGCCTCTCTGTTTCATTT 1421
Db 1141 tgaacctaaacacagaaattgctgcagaaactccttgatttagcctctctgttctatt 1200
QY 1422 CCTCAAAAAGATTCCCCA 1440
Db 1201 cctcaaaaagatttccccca 1219

RESULT 8

AAI63897
ID AAI63897 standard; cDNA; 596 BP.
XX
AC AAI63897;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 105.
XX
KW Human; antiarthritic; antirheumatic; antiproliferative; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; cytostatic; immunosuppressive; nootropic;

KW neuroprotective; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; anticonvulsant; antibacterial;
KW antiparasitic; cardiant; gene therapy; cancer; immune disorder;
XX cardiovascular disorder; neurological disease; infection; human; ss.
OS Homo sapiens.
XX
PN WO200155308-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01309.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0218647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.

PR	21-SEP-2000;	2000US-0234223.	PR	08-DEC-2000;	2000US-0251990.
PR	21-SEP-2000;	2000US-0234274.	PR	11-DEC-2000;	2000US-0254097.
PR	25-SEP-2000;	2000US-0234997.	XX	05-JAN-2001;	2001US-0259678.
PR	25-SEP-2000;	2000US-0234998.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	26-SEP-2000;	2000US-0235484.	XX	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235834.	PI	P-PSDB; RAM43591.	
PR	27-SEP-2000;	2000US-0235836.	XX	WPI; 2001-488781/53.	
PR	29-SEP-2000;	2000US-0236327.	DR	New isolated nucleic acids and polypeptides, useful for diagnosing,	
PR	29-SEP-2000;	2000US-0236367.	PT	treating and/or preventing human diseases and disorders -	
PR	29-SEP-2000;	2000US-0236368.	XX	Claim 1; SEQ ID NO 105; 664pp + Sequence Listing; English.	
PR	29-SEP-2000;	2000US-0236369.	XX	The invention relates to human polynucleotides (AAI63803-AAI64012) and	
PR	29-SEP-2000;	2000US-0236370.	CC	the encoded proteins (AAI634497-AAI63660) useful for preventing, treating	
PR	02-OCT-2000;	2000US-0236802.	CC	or ameliorating medical conditions e.g. by protein or gene therapy. The	
PR	02-OCT-2000;	2000US-0237037.	CC	genes were isolated from a range of human tissues disclosed in the	
PR	02-OCT-2000;	2000US-0237038.	CC	specification. The nucleic acids, proteins, antibodies and (ant)agonists	
PR	02-OCT-2000;	2000US-0237039.	CC	are useful in the diagnosis, treatment and prevention of: (a) cancer,	
PR	12-OCT-2000;	2000US-0237040.	CC	e.g. breast and ovarian cancer and other cancers of the adrenal gland,	
PR	13-OCT-2000;	2000US-0239935.	CC	bone, bone marrow, breast, gastrointestinal tract, liver, lung, or	
PR	13-OCT-2000;	2000US-0239937.	CC	urogenital; (b) immune disorders e.g. Addison's disease, allergies,	
PR	20-OCT-2000;	2000US-0240960.	CC	autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,	
PR	20-OCT-2000;	2000US-0241221.	CC	Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative	
PR	20-OCT-2000;	2000US-0241785.	CC	colitis; (c) cardiovascular disorders such as myocardial ischaemia;	
PR	20-OCT-2000;	2000US-0241786.	CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and	
PR	20-OCT-2000;	2000US-0241808.	CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal	
PR	20-OCT-2000;	2000US-0241809.	CC	and parasitic infections.	
PR	20-OCT-2000;	2000US-0241826.	CC	Note: The sequence data for this patent did not form part of the	
PR	01-NOV-2000;	2000US-0244617.	CC	printed specification, but was obtained in electronic format directly	
PR	08-NOV-2000;	2000US-0246474.	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
PR	08-NOV-2000;	2000US-0246475.	XX	Sequence 596 BP; 133 A; 157 C; 175 G; 126 T; 5 other;	
PR	08-NOV-2000;	2000US-0246476.			
PR	08-NOV-2000;	2000US-0246477.			
PR	08-NOV-2000;	2000US-0246478.			
PR	08-NOV-2000;	2000US-0246523.			
PR	08-NOV-2000;	2000US-0246524.			
PR	08-NOV-2000;	2000US-0246525.			
PR	08-NOV-2000;	2000US-0246526.			
PR	08-NOV-2000;	2000US-0246527.			
PR	08-NOV-2000;	2000US-0246528.			
PR	08-NOV-2000;	2000US-0246532.			
PR	08-NOV-2000;	2000US-0246609.			
PR	08-NOV-2000;	2000US-0246610.			
PR	08-NOV-2000;	2000US-0246611.			
PR	08-NOV-2000;	2000US-0246613.			
PR	17-NOV-2000;	2000US-0249207.			
PR	17-NOV-2000;	2000US-0249208.			
PR	17-NOV-2000;	2000US-0249209.			
PR	17-NOV-2000;	2000US-0249210.			
PR	17-NOV-2000;	2000US-0249211.			
PR	17-NOV-2000;	2000US-0249212.			
PR	17-NOV-2000;	2000US-0249213.			
PR	17-NOV-2000;	2000US-0249214.			
PR	17-NOV-2000;	2000US-0249215.			
PR	17-NOV-2000;	2000US-0249216.			
PR	17-NOV-2000;	2000US-0249217.			
PR	17-NOV-2000;	2000US-0249218.			
PR	17-NOV-2000;	2000US-0249244.			
PR	17-NOV-2000;	2000US-0249245.			
PR	17-NOV-2000;	2000US-0249264.			
PR	17-NOV-2000;	2000US-0249265.			
PR	17-NOV-2000;	2000US-0249297.			
PR	17-NOV-2000;	2000US-0249299.			
PR	17-NOV-2000;	2000US-0249300.			
PR	01-DEC-2000;	2000US-0250160.			
PR	01-DEC-2000;	2000US-0250391.			
PR	05-DEC-2000;	2000US-0251030.			
PR	05-DEC-2000;	2000US-0251988.			
PR	05-DEC-2000;	2000US-0256719.			
PR	06-DEC-2000;	2000US-0251479.			
PR	08-DEC-2000;	2000US-0251856.			
PR	08-DEC-2000;	2000US-0251868.			
PR	08-DEC-2000;	2000US-0251869.			
PR	08-DEC-2000;	2000US-0251989.			

Query Match	38.7%;	Score 557.6;	DB 22;	Length 596;
Best Local Similarity	97.6%;	Pred. No. 7.4e-152;		
Matches 571;	Conservative 4;	Mismatches 9;	Indels 1;	Gaps 1;

QY	166	CCCGCCCCCAAAAGAGCGGAGATTTAACGGGGAGCTGCGGCCAGAGCTGGGGAAATGG	225
Db	13	cccccccccaaaagcggagatttaacgggagctgcccagagctggggaatgg	72
QY	226	GCCCGCAGCCAGCGCGCTTCTCCTCCTGATGCTTTTCAGACCGCGCGTCTTCGAGG	285
Db	73	gcccgagccagcggcgctc	132
QY	286	GGCGTTGCTGGGTTTCACACTCTTCACACTTCTTCATGGTGGCTCAGAGCAGGACC	345
Db	133	ggcgcttgcgtctcacactctctgcaactctctctctctctctctctctctctctctctc	192
QY	346	TTGGTCTTTCTTGTGTTTGAAGCTTTGGGCTAGTGATGATGATGATGATGATGATGATG	405
Db	193	ttggtcttctctgtttgaagctttggtacgtgagatgacccagctgttcgttctatg	252
QY	406	ATGATGAGAGTCGCCGTGTGGAGCCCGCACTCCATCGGTTTCCAGTAGAATTTCAAGCC	465
Db	253	atcatgagagtcgccgtgtggagccccgaactccatgggtttccagtagaattccaagcc	312
QY	466	AGATGTGGTGCAGTCAGTCAGAGTCTGAAAGGGTGGGATCAGATGTTTCACTGTTGACT	525
Db	313	agatgtggctgcagctgagtcagagctctgaaa-ggaggatcacatgcttcaactgttgact	371
QY	526	TCTGGACTATTATGGAATAATCAACACAGCAAGAGGATCCACACCCCTGCGAGGTATCC	585
Db	372	tctggactattatggraaatcacacacacagcaggrgtyccacacccctgcatcatcct	431
QY	586	TGGCGCTGTAATGCAAGAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATG	645
Db	432	ggggctgtgaaatgcaagaagacacagtagccagggtactggaggtactggaggtacgggtag	491

QY 646 GCGAGGACCACCTTGAATTCCTGACACACTGGATTCGAGAGCAGACAGGAGG 705
|||||
Db 492 ggcaggaccacctgaattctgcctcgacacactggattggagcagcagaaccaggg 551
QY 706 CTTGGCCCAACCAAGCTGGAGTGGGAAAGGACAAAGATTTCGGGCCA 750
Db 552 cctggcccacaaactggagtggaaaggcacaagattcgggcca 596

RESULT 9
AAC68427
ID AAC68427 standard; DNA; 10825 BP.
XX
AC AAC68427;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d2 mutation DNA.
XX
DE HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
DR WPI; 2001-006341/01.
DR P-PSDB; AAB36871.
XX
XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;

Query Match 22.3%; Score 321; DB 22; Length 10825;
Best Local Similarity 72.1%; Pred. No. 2.2e-82;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;

QY 298 GTTCACACTCTCTGCACACTCTTCATGGTGGCTCAGACAGGACCTTGGCTTTCCT 357
|||||
Db 3762 gtccacactctctgcacactcttcacatgggtgcctcagcagcagacctgtgttctct 3821
QY 358 TGTTTGAAGCTTTGGGCTACGTTGGATGACCAAGCTTCTCGTGTCTATGATGATGAGATC 417
|||||
Db 3822 ttttgaagctttggctacgtggatgaccagctgttccgtttctctatgatgatgagatc 3881
QY 418 GCGGTGTGAGCCCGCACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
|||||

Db 3882 gccgtgtgagcccccgaactccatggtttccagtagaattccaagcagatgtgctgc 3941
QY 478 AGCTGAGTCAGACTCTGAAAGGGTGGGATCACATGTTCACTGTTCACTTCTGGACTATTA 537
|||||
Db 3942 agctgagtcagagctctgaaagggtggatcacatgttcaactgttgactctctggactatta 4001
QY 538 TGGAAATTCACAACACAGCAAG----- 560
|||||
Db 4002 tggaaatcacacacacagcaagggtatgtgagagggggcctcaccttctcctgaggtgtg 4061
QY 561 ----- 560
Db 4062 cagagcttttcatcttttcatgcatcttgaaggaaacacagctgggaagtgtgaggtctgtg 4121
QY 561 ----- 560
Db 4122 ggagcaggaagaggaaggaattgtctctcctgagatcatttgccttggggatgtg 4181
QY 561 -----GAGTCCCA 568
|||||
Db 4182 aaatagggaacctattctcttggttgagtttaacaagctggggatttttccagagttccca 4241
QY 569 CACCCTGCAGGTCACTCTGGGCTGTGAATGCAAGAGACAAACAGTACCAGGCGCTACTG 628
|||||
Db 4242 caccctcaggtcatcctcctgggtgtgaaatgcaagagacacagtcacagtgaggtactg 4301
QY 629 GAAGTACGGGTATGATGGGAGGACACACCTTGAAATTCCTGCCCTGCACACTGGATTGGAG 688
|||||
Db 4302 gaagtacgggtatgtatgggcaggaaccaccttgaaattgccttgacacactggtatggag 4361
QY 689 AGCAGCAGAACCCAGGCGCTGGGCCACCAAGCTGGAGTGGGAAAGGCACAGATTCCGGC 748
|||||
Db 4362 agcagcagaaccaggcctggccaccacagctggagtggaaaggcacaagattcgggc 4421
QY 749 CAGGCAGAACAGGCGCTTACCTGGAGAGGGAGTGCCTGCACAGCTGCAGCAGTTGCTGGA 808
Db 4422 caggcagaaacaggccaccctggagaggactgccctgcacagctgcagcagttgctgga 4481
QY 809 GCTGGGGAGAGGTGTTTGGACCAACAAG 837
|||||
Db 4482 gctggggagaggtgttttggaccaacaag 4510

RESULT 10
AAC68428
ID AAC68428 standard; DNA; 10825 BP.
XX
AC AAC68428;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
DE HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
OS Homo sapiens.
XX
PN US6140305-A.
XX
PD 31-OCT-2000.
XX
PF 04-APR-1997; 97US-0834497.
XX
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
PR 23-MAY-1996; 96US-0652265.
XX
PA (BIRA) BIO-RAD LAB INC.
PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX

DR WPI; 2001-006341/01.
XX P-PSDB; AAB36872.
PT New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
XX chelation agent alleviating iron overload -
XX
PS Disclosure; Fig 3; 108pp; English.
XX
CC The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2252 C; 2548 G; 2926 T; 0 other;

Query Match 22.3%; Score 321; DB 22; Length 10825;
Best Local Similarity 72.1%; Pred No. 2, 2e-82;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;

QY 298 GTTCACACTCTCTGCACTACCTCTTCATGGTGCGTCCCTCAGAGCAGGACCTTGCTCTTCT 357
Db |||||||
QY 3762 gtccacactctctgcactacctctctcatggtgctccagagcaggacctggtctcttct 3821
Db |||||||
QY 358 TGTTTGAAGCTTTGGCGTACGTGGGATGACCAGCTGTCTGCTTCTATGATGATGAGATC 417
Db |||||||
QY 3822 tgtttgaagctttgggtcgtggtgaccagctgtctgtctatgatgatgagagtc 3881
Db |||||||
QY 418 GCCGTGTGGAGCCCCGAATCCATGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
Db |||||||
QY 3882 gccgtgtggagcccccgaactccatggtttccagtagaatttcaagccagatgtggtctc 3941
Db |||||||
QY 478 AGCTGAGTCAGACTCTGAAGGGTGGGATCACATGTTCTACTGTTGACTTCTGGACTATT 537
Db |||||||
QY 3942 agctgagtcagagtcgaaagggtgggtgacatgttctactgttgcactctggactatta 4001
Db |||||||
QY 538 TGGAAATCACACACACACCAAG----- 560
Db tggaaaatcacacacacagcagggttatgtggagagggggccctcacttcctcctgaggtgtg 4061
QY 561 ----- 560
Db cagagcttttcatcttttcatcatcttgaaagaaacagctggaagtctgaggtctgtg 4121
QY 561 ----- 560
Db ggagcaggggaagaggaaggaattgtctctcctgagatcatttggctcctggggatggtg 4181
QY 561 -----GAGTCCCA 568
Db aaatagggacctattcctcttgggtcagttatacaagctgaggattttccagagtcacca 4241
QY 569 CACCCTGCAGGTCTATCCTGGGCTGTGAATGCAAGAGAACACAGTACCCAGGGCTACTG 628
Db |||||||
QY 4242 caccctgcaggtcatctcctggctgtgaaatgcaagaagacaacagtaaccagggctactg 4301
QY 629 GAAGTACGGGTATGATGGGAGCAGCACCTTGAATTCCTGCCCTGACACACTGGATTGGAG 688
Db |||||||
QY 4302 gaagtaacgggtatgatgggagcagcacccttgaaattctgacctgacacactggattggag 4361
QY 689 AGCAGCAGAACCCAGGGCCCTGGCCCAACCAAGCTGGAGTGGGAAAGGCACAAATTCGGGC 748
Db |||||||
QY 4362 agcagcagaaccacagggccctggcccacaaagctggagtgggaaagcacaagattcgggc 4421
QY 749 CAGCAGAACAGGGCCCTTACTGTGAGAGGGGACTGCCCTGCACAGCTGCAGCAGTTGCTGA 808
Db |||||||
QY 4422 caggcagaacagggccctaccttgagaggggactgcccctgcacagctgcagcagttgtgga 4481
QY 809 GCTGGGAGAGGTCTTTTGGACCAACAG 837
Db |||||||

Db 4482 gctggggagaggtgttttggaccaacaag 4510

RESULT 11
AAT96690
ID AAT96690 standard; DNA; 10825 BP.
XX
XX AAT96690;
XX
XX 14-APR-1998 (first entry)
XX Hereditary haemochromatosis gene.
DE Hereditary haemochromatosis gene.
XX Hereditary haemochromatosis; metal toxicity; diagnosis;
KW gene therapy; prenatal screening; human; ds.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 361..7147
FT /tag= a
FT /note= "contains introns"
FT intron 437..3761
FT /tag= b
FT intron 4026..4234
FT /tag= c
FT intron 4511..5605
FT /tag= d
FT intron 5882..6039
FT /tag= e
FT intron 6154..7106
FT /tag= f
FT mutation 3872
FT /tag= g
FT /note= "C to G substitution (24d2 mutation)
FT results in His to Asp substitution"
FT variation 3878
FT /tag= h
FT /note= "A to T substitution (24d7 variant)
FT results in Ser to Cys substitution"
FT mutation 5834
FT /tag= i
FT /note= "G to A substitution (24dl mutation
FT associated with HH), results in Cys to
FT Tyr substitution"
XX WO9738137-A1.
XX
XX 16-OCT-1997.
PD
XX
XX 04-APR-1997; 97WO-US06254.
PF
XX 23-MAY-1996; 96US-0652265.
PR 04-APR-1996; 96US-0630912.
PR 16-APR-1996; 96US-0632673.
XX (MERC-) MERCATOR GENETICS INC.
PA
XX Drayna DT, Feder JN, Gnirke A, Ruddy D, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
XX WPI; 1997-512743/47.
DR P-PSDB; AAW36499.
XX
XX Hereditary haemochromatosis gene and variants - useful for diagnosis
PT and treatment of hereditary haemochromatosis disease
XX
XX Disclosure; Fig 3; 115pp; English.
PS

[illegible]

RESULT	13	
AAC68426		
ID	AAC68426	standard; DNA; 10825 BP.
XX	XX	
XX	AAC68426;	
XX		
DT	21-FEB-2001	(first entry)
XX		
DE	Human hereditary hemochromatosis 24dl	mutation DNA.
XX		
KW	HH;	hereditary hemochromatosis; chelation agent;
KW	T-cell	differentiation factor; iron overload; ds.
XX		
XX	Homo sapiens.	
XX		
PN	US614,305-A.	
XX		

PD	31-OCT-2000.	
XX		
XX	04-APR-1997;	97US-0834497.
XX		
XX	04-APR-1996;	96US-0630912.
PR	16-APR-1996;	96US-0632673.
PR	23-MAY-1996;	96US-0652265.
XX		
XX	(BIRA) BIO-RAD LAB INC.	
PA		
XX	Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi 2, Wolff RK;	
PI	Feder JN;	
PI		
XX	WPI; 2001-006341/01.	
DR	P-PSDB; AAB36870.	
DR		
XX		
XX	New hereditary hemochromatosis gene products or polypeptides, useful	
PT	for treating hereditary hemochromatosis in a patient, and as a metal	
PT	chelation agent alleviating iron overload -	
XX		
XX	Disclosure; Fig 3; 10pp; English.	
XX		
CC	The present invention relates to hereditary hemochromatosis gene	
CC	products. These proteins may be used to treat a patient diagnosed as	
CC	having human hemochromatosis disease. It is also useful as a metal	
CC	chelation agent or as a T-cell differentiation factor, and for	
CC	alleviating iron overload. They may also be used in protein replacement	
CC	therapy for individuals having a defective human hemochromatosis gene.	
XX		
XX	Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;	
SQ		
	Query Match 22.2%; Score 319.4; DB 22; Length 10825;	
	Best Local Similarity 72.08; Pred. No. 6.4e-82;	
	Matches 533; Conservative 0; Mismatches 1; Indels 209; Gaps	
Qy	298 GTTCACACTCTCTGCACCTCTCTTCATGCGTCTCAGACGAGACCTTGGTCTTTCTT	357
Db		
Db	3762 gtccacactctcgactcactctcttcattgggtcctcagcagcagccttgctcttcc	3821
Qy	358 TGTGTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCTGCTTCTATGATGAGAGTC	417
Db		
Db	3822 tgtttgaagcttttgggtcactgtgatgacacagctgttcgtttctatgatcatgagagtc	3881
Qy	418 GCCGTGTGGAGCCCGCAACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGGCGTGC	477
Db		
Db	3882 gccgtgtgagcccgcaactccatgggtttccagtagaattccaagccagatgtggctgc	3941
Qy	478 AGCTGAGTCAGAGCTGAAGGGTGGGATCACATGTTCTACCTGTTCTGAGCTATTATTA	537
Db		
Db	3942 agctgagtcagtgctgaaagggtgggtacacatgtctcactgttgactctctggactatta	4001
Qy	538 TGGAAAAATCACAAACCACACAGCAAG-	560
Db		
Db	4002 tggaaaaatcacaaaccacagcaagggtatgtggagagggggccctcccttccctgaggtt	4061
Qy	561	560
Db	4062 cagagcttttcatcttttcatgcatcttgaaggaaacagctggaagctctgaggtcttgc	4121
Qy	561	560
Db	4122 ggagcagggaaagaggaaatttgccttcctgagatcatttggcttccttggggatggttgc	4181
Qy	561	568
Db	4182 aaatagggaacctattcctttgttgcagttaacaaggctggggatttttccagagatccca	4241
Qy	569 CACCCTGAGGTCATCTGGGCTGTGAATGCAAGAAGACACAGTACCGAGGGCTACTTG	628
Db		
Db	4242 caccctcgagggtca tccctggggtgtgaaatgcaagaagacacacagtagtaccagggctactg	4301
Qy	629 GAAGTACGGGTATGATGGCGAGGACCACTTGAAATTCTGCCCTGACACACTGGATGGAG	688

Db 4302 gaagtacgggtatgaggcagagaccactgaattctgcctgcacactggattggag 4361
QY 689 ACAGCAGAAACCCAGGCGCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAGATTGCGGC 748
Db 4362 agcagcagaaccaggcctgcccaccaggctggagtggaaaggcacaagattcgggc 4421
QY 749 CAGGCGAAGACAGGCGCTACTCTGAGAGGAGCTGCCCTGCACAGCTGCACAGTGTGCTGGA 808
Db 4422 caggcagaacaggcctacctgagaggactgccctgcacagctgcagcagttgctgga 4481
QY 809 GCTGGGAGAGGTGTTTGGACCAACAAG 837
Db 4482 gctggggagaggtgttttggaccacaag 4510
RESULT 14
AAA96794
ID AAA96794 standard; cDNA; 12146 BP.
XX AC AAA96794;
XX DT 19-FEB-2001 (first entry)
XX DE Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX KW Human; histocompatibility iron loading protein; HFE protein;
KW major histocompatibility complex; non-classical class I gene;
KW chromosome 6p; iron disorder; haemochromatosis; ss.
OS Homo sapiens.
XX FH Key Location/Qualifiers
FT exon 1028..1324
FT FT /*tag= a
FT /*number= 1
FT intron 1325..4651
FT FT /*tag= b
FT /*number= 1
FT exon 4652..4915
FT FT /*tag= c
FT /*number= 2
FT intron 4916..5124
FT FT /*tag= d
FT /*number= 2
FT exon 5125..5400
FT FT /*tag= e
FT /*number= 3
FT intron 5401..6493
FT FT /*tag= f
FT /*number= 3
FT exon 6494..6769
FT FT /*tag= g
FT /*number= 4
FT intron 6770..6927
FT FT /*tag= h
FT /*number= 4
FT exon 6928..7041
FT FT /*tag= i
FT /*number= 5
FT intron 7042..7994
FT FT /*tag= j
FT /*number= 5
FT exon 7995..9050
FT FT /*tag= k
FT /*number= 6
FT intron 9051..10205
FT FT /*tag= l
FT /*number= 6
FT exon 10206..10637
FT FT /*tag= m
XX PN WO200058515-A1.

XX 05-OCT-2000.
PD 24-MAR-2000; 2000WO-US07982.
XX 26-MAR-1999; 99US-0277457.
PR (BILL-) BILLUPS-ROTHENBERG INC.
XX Rothenberg BE, Sawada-Hirai R, Barton JC;
PI WPI; 2000-647244/62.
XX Diagnosing an iron disorder e.g. hemochromatosis or a genetic
PT susceptibility to develop it, by determining the presence of a mutation
PT in exon 2 or an intron of a histocompatibility iron loading nucleic
PT acid -
XX Example 1; Page 21-28; 55pp; English.
XX The present sequence represents the human histocompatibility iron
CC loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
CC non-classical class I gene located on chromosome 6p. Mutations in the
CC gene lead to iron disorders. The specification describes a method for
CC diagnosing an iron disorder or a genetic susceptibility to develop the
CC disorder in a mammal. The method comprises determining the presence of
CC a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
CC is not a C to G missense mutation at nucleotide 187 of the sequence
CC given in A96769 (Genbank Accession number U60319). The presence of the
CC mutation indicates the disorder or the genetic susceptibility to the
CC disorder. The method is used to diagnose an iron disorder
CC e.g. haemochromatosis, or a genetic susceptibility to develop it.
XX SQ Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;
Query Match 22.2%; Score 319.4; DB 21; Length 12146;
Best Local Similarity 72.0%; Pred. No. 6.8e-82;
Matches 539; Conservative 0; Mismatches 1; Indels 209; Gaps 1;
QY 298 GTTCACACTCTCTGCACACTACCTCTTCATGGTGCCTCAGACAGGAGCTTGGCTTTCCCT 357
Db 4652 gtccacactctctgcactacctcttcattgggtgcctcagagcaggaccttggcttctcct 4711
QY 358 TGTTTGAAGCTTTGGGCTACGTGGATGACCATGCTGTTCGTGTTCTATGATGATGAGAGTC 417
Db 4712 tgttgaagctttgggctacgtggatgaccagctgttcgtgttctatgcatgagagtc 4771
QY 418 GCGGTGTGGAGCCCCGAACTCCATGGTTTCCAGTAGAATTCAGCCACAGATGGCTGC 477
Db 4772 gccgtgtgagccccgaactccatgggtttccagtagaatttcaagccagatgtggctgc 4831
QY 478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTA 537
Db 4832 agctgagtcagagctgaaagggtggatcacatgttccactgtgacttctggaactatta 4891
QY 538 TGGAAAATCACAAACCACAGCAAG----- 560
Db 4892 tggaaaatcacaaaccacagcaagggtatgtggagagggggccctcaccttctcagggtgtg 4951
QY 561 ----- 560
Db 4952 cagagcttttcatcttttcattcatcttgaggaagaacagctggaagtctgaggtctgtg 5011
QY 561 ----- 560
Db 5012 ggagcaggggaagagggaagaatttgccttcctgagatcatttggctccttgggatggtg 5071
QY 561 -----GAGTCCCA 568
Db 5072 aaatagggaacctattcctttgttggttcagttacaaggctggggtattttccagagtccca 5131
QY 569 CACCCCTGCAGGTCACTCCTGGGCTGTGAAATGCAAGAAGACAACAGTACCGAGGGCTACTG 628

|||||
Db 5132 caccctgcagtcatectggctgtaaatgcaagaagacaacagtcacgggctactg 5191
QY 629 GAAGTACGGGTATGATGGCGCAGGACCACTTGAATCTGCCCTGCACACACTGGATTGGAG 688
Db 5192 gaagtacgggtatgatggcagagcacacttgaaattctgacctgcacacactggattggag 5251
QY 689 AGCAGCAGAACCCAGGGCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAAGATTTCGGGC 748
Db 5252 agcagcagaacccagggcctggccaccacaagctggagtg99aaaggcaagattcgggc 5311
QY 749 CAGGCAGAACAGGGCTACCTGGAGAGGGAGTGCCTGCACACAGCTGCACAGTTCCTGCA 808
Db 5312 caggcagaacaggcctacctgagaggactgcctgcacagctgcagcagttgctgga 5371
QY 809 GCTGGGAGAGGTGTTTGGACCAACAG 837
Db 5372 gctggggagaggtgttttggaccacaag 5400
RESULT 15
AAV57903/C
ID AAV57903 standard; DNA; 237326 BP.
XX
AC AAV57903;
XX
DT 21-DEC-1998 (first entry)
DE
XX Hereditary haemochromatosis subregion from an HH affected individual.
XX
XX Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;
KW diagnosis; iron metabolism; NPT3; NPT4; Roret; BTF1; BTF2; BTF3;
KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;
KW type 1 sodium transport gene; ss.
XX
OS Homo sapiens.
XX
XX WO9814466-A1.
XX
PD 09-APR-1998.
XX
PF 30-SEP-1997; 97WO-US17658.
XX
PR 07-MAY-1997; 97US-0852495.
PR 01-OCT-1996; 96US-0724394.
XX
PA (PROG-) PROGENITOR INC.
XX
XX Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;
PI Tsuchihashi Z, Wolff RK;
XX
XX WPI; 1998-240014/21.
DR
PT Hereditary haemochromatosis gene products - used to develop products
PT for the diagnosis and treatment of hereditary disorders in iron
PT metabolism
XX
PS Claim 1; Fig 9; 209pp; English.
XX
CC The present invention describes hereditary haemochromatosis gene
CC products from the human haemochromatosis gene. The present sequence
CC represents a hereditary haemochromatosis subregion from an hereditary
CC haemochromatosis (HH) affected individual. Also described is a
CC method to determine the presence or absence of the common hereditary
CC haemochromatosis (HFE) gene mutation in an individual comprising:
CC (a) providing DNA or RNA from the individual; and (b) assessing the
CC DNA or RNA for the presence or absence of a haplotype or genotype where
CC the presence or absence of the haplotype genotype indicates the likely
CC presence of the HFE gene mutation in the genome of the individual. The
CC HFE gene sequences from the present invention can be used to develop
CC products for use in the diagnosis and treatment of HFE. The present
CC invention also describes BTF genes, which are homologues of the milk
CC protein, butyrophilin (BT), and can be used in the production of agonists

CC and antagonists of BT function. Also described are: (1) a Roret gene
CC which can be used to develop products for the study, diagnosis and
CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes
CC which are homologues of a type 1 sodium transport gene, and can
CC similarly be used for hypophosphatemia.
XX
SQ Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

Query Match 22.2%; Score 319.4; DB 19; Length 237326;
Best Local Similarity 72.0%; Pred. No. 3e-81;
Matches 539; Conservative 0; Mismatches 1; Indels 209; Gaps 1;

QY 298 GTTTCACACTCTCTGCACCTACCTCTTCAATGGGTGCCTCAGAGCAGGACCTTGGCTTTCCT 357
Db 43338 GTTTCACACTCTCTGCACCTACCTCTTCAATGGGTGCCTCAGAGCAGGACCTTGGCTTTCCT 43279
QY 358 TGTTTGAAGCTTTGGGCTACGTGGATGCCAGCTGTTTCGTTCTATGATGATGAGAGTC 417
Db 43278 TGTTTGAAGCTTTGGGCTACGTGGATGCCAGCTGTTTCGTTCTATGATGATGAGAGTC 43219
QY 418 GCCGTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
Db 43218 GCCGTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 43159
QY 478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTGA 537
Db 43158 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTGA 43099
QY 538 TGGAAATCAACACCAAGCAAG----- 560
Db 43098 TGGAAATCAACACCAAGCAAGGTATGTGGAGAGGGGGCCCTCACCCTTCTGAGGTTGT 43039
QY 561 ----- 560
Db 43038 CAGAGCTTTTCATCTTTTCATGTCATCTTGAAGGAACAGCTGGAAGTCTGAGGCTTGTG 42979
QY 561 ----- 560
Db 42978 GGAGCAGGAAGAGGAAGGAATTTGCTTCTGAGATCATTTGGTCTCTGGGATGGTG 42919
QY 561 -----CAGTCCCA 568
Db 42918 AATAGGGACCTATTCCTTTGGTTGGAGTTAACAAAGCTGGGATTTTCCAGAGTCCCA 42859
QY 569 CACCTGCAGGTCTCTCTGGGCTGTGAAATGCAAGAGAACACAGTACCGAGGGCTACTG 628
Db 42858 CACCTGCAGGTCTCTCTGGGCTGTGAAATGCAAGAGAACACAGTACCGAGGGCTACTG 42799
QY 629 GAAGTACGGGTATGATGGGCGAGGACCACTTGAATTTGCCCCTGACACACTGGATTGGAG 688
Db 42798 GAAGTACGGGTATGATGGGCGAGGACCACTTGAATTTGCCCCTGACACACTGGATTGGAG 42739
QY 689 AGCAGCAGAACCCAGGGCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAAGATTTCGGGC 748
Db 42738 AGCAGCAGAACCCAGGGCTGGCCACCAAGCTGGAGTGGGAAAGGCACAAAGATTTCGGGC 42679
QY 749 CAGCAGAACAGGGCTTACCTGGAGAGGGAGTGCCTCTGCACAGCTGCAGCAGTTCGTGA 808
Db 42678 CAGCAGAACAGGGCTTACCTGGAGAGGGAGTGCCTCTGCACAGCTGCAGCAGTTCGTGA 42619
QY 809 GCTGGGAGAGGTGTTTGGACCAACAG 837
Db 42618 GCTGGGAGAGGTGTTTGGACCAACAG 42590

Search completed: June 19, 2002, 08:53:55
Job time: 8120 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:40:03 ; Search time 79.08 Seconds
(without alignments)
4472.841 Million cell updates/sec

Title: US-09-497-957-12

Perfect score: 1440

Sequence: 1 GGGGACACTGGATCAGCTAG.....TCTCAAAAAGATTTCCCA 1440

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	100.0	1440	3	US-08-652-265-12
2	1440	100.0	1440	3	US-08-834-497A-12
3	1440	100.0	1440	4	US-09-503-444A-12
4	1438.4	99.9	1440	3	US-08-652-265-10
5	1438.4	99.9	1440	3	US-08-652-265-11
6	1438.4	99.9	1440	3	US-08-834-497A-10
7	1438.4	99.9	1440	3	US-08-834-497A-11
8	1438.4	99.9	1440	4	US-09-503-444A-10
9	1438.4	99.9	1440	4	US-09-503-444A-11
10	1436.8	99.8	1440	3	US-08-652-265-9
11	1436.8	99.8	1440	3	US-08-834-497A-9
12	1436.8	99.8	1440	4	US-09-503-444A-9
13	1215.8	84.4	2506	4	US-09-277-457-1
14	321	22.3	10825	3	US-08-652-265-5
15	321	22.3	10825	3	US-08-652-265-7
16	321	22.3	10825	3	US-08-834-497A-5
17	321	22.3	10825	3	US-08-834-497A-7
18	321	22.3	10825	4	US-09-503-444A-7
19	321	22.3	10825	4	US-09-503-444A-5
20	319.4	22.2	10825	3	US-08-652-265-1
21	319.4	22.2	10825	3	US-08-652-265-3
22	319.4	22.2	10825	3	US-08-834-497A-1
23	319.4	22.2	10825	3	US-08-834-497A-3
24	319.4	22.2	10825	4	US-09-503-444A-1
25	319.4	22.2	10825	4	US-09-503-444A-3
26	319.4	22.2	12146	4	US-09-277-457-27
27	319.4	22.2	246240	2	US-08-724-394A-20

28 319.4 22.2 246240 2 US-08-724-394A-21 Sequence 21, Appl
29 319.4 22.2 246240 2 US-08-724-394A-22 Sequence 22, Appl
30 280 19.4 517 1 US-08-632-673B-4 Sequence 4, Appl
31 280 19.4 517 3 US-08-652-265-21 Sequence 21, Appl
32 280 19.4 517 3 US-08-834-497A-21 Sequence 21, Appl
33 280 19.4 517 4 US-09-503-444A-21 Sequence 21, Appl
34 279 19.4 517 1 US-08-632-673B-13 Sequence 13, Appl
35 278.4 19.3 360 3 US-08-905-124-5 Sequence 5, Appl
36 278.4 19.3 517 1 US-08-632-673B-3 Sequence 3, Appl
37 278.4 19.3 517 3 US-08-652-265-20 Sequence 20, Appl
38 278.4 19.3 517 3 US-08-834-497A-20 Sequence 20, Appl
39 278.4 19.3 517 4 US-09-503-444A-20 Sequence 20, Appl
40 173.6 12.1 1112 3 US-08-890-719-5 Sequence 5, Appl
41 172 11.9 1230 3 US-08-890-719-9 Sequence 10, Appl
42 168.8 11.7 1195 3 US-08-890-719-10 Sequence 7, Appl
43 166.8 11.6 1197 3 US-08-890-719-7 Sequence 37, Appl
44 166.8 11.6 1262 3 US-08-890-719-37 Sequence 4, Appl
45 162 11.2 1145 3 US-08-890-719-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-08-652-265-12
; Sequence 12, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(408, "g")

OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d1
US-08-652-265-12

Query Match 100.0%; Score 1440; DB 3; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGACATGGATCAGTACCTGTTTCAACAAGCAGTACCTCTGCTGTAGGAGAGAGA 60
Db 1 GGGGACATGGATCAGTACCTGTTTCAACAAGCAGTACCTCTGCTGTAGGAGAGAGA 60

Qy 61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCTGAAAGACCTGTTGCTTTTACCAGGAAGTTTACTGGGCATCTCCTGAG 120

Qy 121 CCTAGCAATAGCTGTAGGTCAGTCTGCGAGCCATCCCGTTTCCCGCCGCCCAAG 180
Db 121 CCTAGCAATAGCTGTAGGTCAGTCTGCGAGCCATCCCGTTTCCCGCCGCCCAAG 180

Qy 181 AAGCGAGATTTAAAGCGGACCTGCGGCGAGAGCTGGGAAATGGCGCGGAGCAGGC 240
Db 181 AAGCGAGATTTAAAGCGGACCTGCGGCGAGAGCTGGGAAATGGCGCGGAGCAGGC 240

Qy 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTGCGAGGGCGCTTGTGCGTT 300
Db 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTGCGAGGGCGCTTGTGCGTT 300

Qy 301 CACACTCTGCACTACCTCTTCAATGGGTGCTCAGACAGACCTTGGTCTTCTTGT 360
Db 301 CACACTCTGCACTACCTCTTCAATGGGTGCTCAGACAGACCTTGGTCTTCTTGT 360

Qy 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGCTGCTGTTCTATGATGATGAGTCCGC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGCTGCTGTTCTATGATGATGAGTCCGC 420

Qy 421 GTGTGAGCCCGCACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGAGCCCGCACTCCATCGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480

Qy 481 TGAGTCAGAGTCTGAAGGGTGGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCAGATGTTCACTGTTGACTTCTGGACTATTATGG 540

Qy 541 AAAATCACAAACACAGCAAGGAGTCCACACCTGTCAGGTCTATCTGGGCTGTGAATGC 600
Db 541 AAAATCACAAACACAGCAAGGAGTCCACACCTGTCAGGTCTATCTGGGCTGTGAATGC 600

Qy 601 AAGAAGACAACAGTACCGAGGCTTACTGGAAGTACGGTATGATGGCAGGACCACTTG 660
Db 601 AAGAAGACAACAGTACCGAGGCTTACTGGAAGTACGGTATGATGGCAGGACCACTTG 660

Qy 661 AATCTGCGCTTACACACTTGGATGGAGAGCAGACCAAGCCAGGCGCTGGCCCAACAGC 720
Db 661 AATCTGCGCTTACACACTTGGATGGAGAGCAGACCAAGCCAGGCGCTGGCCCAACAGC 720

Qy 721 TGGAGTGGGAAAGGCAACAAGATTCGGGCGAGGAGTCCAGAGGCGCTTCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAAGGCAACAAGATTCGGGCGAGGAGTCCAGAGGCGCTTCTGGAGAGGACT 780

Qy 781 GCCTTGCACAGCTGCAGCAGTGTCTGGAGCTGGGGAGAGGTTTGGACCAACAAGTGC 840
Db 781 GCCTTGCACAGCTGCAGCAGTGTCTGGAGCTGGGGAGAGGTTTGGACCAACAAGTGC 840

Qy 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGAGTACCACTCTACGGTGTGGG 900

Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTCTACGGTGTGGG 900

Qy 901 CCTTGAAGTACTACCCCAAGACATCACTGAAGTGGCTGAAGGATGAAGCAATGG 960

Db 901 CCTTGAAGTACTACCCCAAGACATCACTGAAGTGGCTGAAGGATGAAGCAATGG 960

Qy 961 ATGCCAAGGAGTTCGAACCTTAAAGACGTAATGCCCCAATGGGATGGGACCTACCAAGGCT 1020

Db 961 ATGCCAAGGAGTTCGAACCTTAAAGACGTAATGCCCCAATGGGATGGGACCTACCAAGGCT 1020

Qy 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGCAGACATATACGTACCAAGTGGAGCACC 1080

Db 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGCAGACATATACGTACCAAGTGGAGCACC 1080

Qy 1081 CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Db 1081 CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

Qy 1141 TTGGAGTCATCAGTGAATTTGCTTTTGTGTCATCTTGTGTCATCTTGTGTCATTTGTTTCA 1200

Db 1141 TTGGAGTCATCAGTGAATTTGCTTTTGTGTCATCTTGTGTCATCTTGTGTCATTTGTTTCA 1200

Qy 1201 TAATATTAAAGAGAGCAGGCTTCAAGAGAGCCATGGGAGGAGCACTAGCTCTTAGCTGAAC 1260

Db 1201 TAATATTAAAGAGAGCAGGCTTCAAGAGAGCCATGGGAGGAGCACTAGCTCTTAGCTGAAC 1260

Qy 1261 GTGAGTGACAGCAGCCTGCAGACTCACTGTGGGAAGGAGAGCACTAGAGACTCAAG 1320

Db 1261 GTGAGTGACAGCAGCCTGCAGACTCACTGTGGGAAGGAGAGCACTAGAGACTCAAG 1320

Qy 1321 AGGAGTGATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Db 1321 AGGAGTGATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380

Qy 1381 TGCTGACGAACTCCTTGATTTAGCCTTCTGTTTCAATTTCTCAAAAAGATTTCCCA 1440

Db 1381 TGCTGACGAACTCCTTGATTTAGCCTTCTGTTTCAATTTCTCAAAAAGATTTCCCA 1440

RESULT 2
US-08-834-497A-12
; Sequence 12, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/632, 673
FILING DATE: 16-APR-1996
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630, 912
FILING DATE: 04-APR-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0056-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d1

US-08-834-497A-12

Query Match 100.0%; Score 1440; DB 3; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGGACACTGGATCACCCTAGTGTTCACAAGCAGGACCTTCTGCTGTAGGAGAGAGA	60
Db	1	GGGACACTGGATCACCCTAGTGTTCACAAGCAGGACCTTCTGCTGTAGGAGAGAGA	60
QY	61	ACTAAAGTCTGAAAGACCTGTGCTTTTCCAGGAAGTTTACTGGGCATCTCTGAG	120
Db	61	ACTAAAGTCTGAAAGACCTGTGCTTTTCCAGGAAGTTTACTGGGCATCTCTGAG	120
QY	121	CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCCGCCCAAAAG	180
Db	121	CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCCGCCCAAAAG	180
QY	181	AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATTTGGCCCGCGAGCCAGGC	240
Db	181	AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATTTGGCCCGCGAGCCAGGC	240
QY	241	CGCGCTTCTCCCTGATGCTTTTCCAGACCGCGTCTGCGAGGGCGCTTGTCTCGTT	300
Db	241	CGCGCTTCTCCCTGATGCTTTTCCAGACCGCGTCTGCGAGGGCGCTTGTCTCGTT	300
QY	301	CACACTCTGCACTACCTCTTCATGGGTGCTTCAGAGCAGGACCTTGGTCTTTCCCTTGT	360
Db	301	CACACTCTGCACTACCTCTTCATGGGTGCTTCAGAGCAGGACCTTGGTCTTTCCCTTGT	360
QY	361	TTGAAGCTTTGGGCTACGTGGATGACACGTGTTCTGTTCTATGATGATGAGTCGCC	420
Db	361	TTGAAGCTTTGGGCTACGTGGATGACACGTGTTCTGTTCTATGATGATGAGTCGCC	420

QY	421	GTGTGGAGCCCGAACTCCATCGTGTTCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC	480
Db	421	GTGTGGAGCCCGAACTCCATCGTGTTCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC	480
QY	481	TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG	540
Db	481	TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG	540
QY	541	AAATCACAACACACAGCAAGGAGTCCCACACCCTGCAGGTCACTCTGGCTGTGAATGC	600
Db	541	AAATCACAACACACAGCAAGGAGTCCCACACCCTGCAGGTCACTCTGGCTGTGAATGC	600
QY	601	AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG	660
Db	601	AAGAAGACAACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG	660
QY	661	AATTCGCCCTGACACACTGGATTTGGAGAGCAGACCCAGGGCCTGGCCCAACAGC	720
Db	661	AATTCGCCCTGACACACTGGATTTGGAGAGCAGACCCAGGGCCTGGCCCAACAGC	720
QY	721	TGGAGTGGGAAAGGCACAGATTCGGGCCAGGCAGACAGGGCCTACCTGGAGAGGACT	780
Db	721	TGGAGTGGGAAAGGCACAGATTCGGGCCAGGCAGACAGGGCCTACCTGGAGAGGACT	780
QY	781	GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTTTGGACCAACAAGTGC	840
Db	781	GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTTTGGACCAACAAGTGC	840
QY	841	CTCCTTTGGTGAAGTGACATCATGTGACCTCTTCACTGACCACTCTACGGTCTCGGG	900
Db	841	CTCCTTTGGTGAAGTGACATCATGTGACCTCTTCACTGACCACTCTACGGTCTCGGG	900
QY	901	CCTTGAACCTACTACCCCAAGACATCACCATGAAGTGGCTGAAGATTAAGCAGCAATGG	960
Db	901	CCTTGAACCTACTACCCCAAGACATCACCATGAAGTGGCTGAAGATTAAGCAGCAATGG	960
QY	961	ATGCCAAGGAGTTGGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAAGGCT	1020
Db	961	ATGCCAAGGAGTTGGAACCTAAAGACGTATTGCCCAATGGGGATGGGACCTACCAAGGCT	1020
QY	1021	GGATAACCTTTGGCTGTACCCCTGGGGAAGACAGATATACGTACCAGGTGGAGCAC	1080
Db	1021	GGATAACCTTTGGCTGTACCCCTGGGGAAGACAGATATACGTACCAGGTGGAGCAC	1080
QY	1081	CAGGCTGGATCAGCCCTCATTTGTGATCTGGAGCCCTCACCGTCTGGCACCCTAGTCA	1140
Db	1081	CAGGCTGGATCAGCCCTCATTTGTGATCTGGAGCCCTCACCGTCTGGCACCCTAGTCA	1140
QY	1141	TTGGAGTCACTAGTGGAAATTTGCTGTTTTTGTGCTCATCTTGTTCATTTGGAAATTTGTTCA	1200
Db	1141	TTGGAGTCACTAGTGGAAATTTGCTGTTTTTGTGCTCATCTTGTTCATTTGGAAATTTGTTCA	1200
QY	1201	TAATATTAAAGGAAGGAGGAGGTTTCAAGAGGAGCCATGGGACCTACCGTCTGGCACCCTAGTCA	1260
Db	1201	TAATATTAAAGGAAGGAGGAGGTTTCAAGAGGAGCCATGGGACCTACCGTCTGGCACCCTAGTCA	1260
QY	1261	GTGAGTACACGACGCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAAG	1320
Db	1261	GTGAGTACACGACGCTGCAGACTCACTGTGGGAAGGAGACAAACTAGAGACTCAAAG	1320
QY	1321	AGGGAGTGCATTTATGAGCTCTTCATGTTTTTTCAGGAGAGTTGAACCTAAACATAGAAAT	1380
Db	1321	AGGGAGTGCATTTATGAGCTCTTCATGTTTTTTCAGGAGAGTTGAACCTAAACATAGAAAT	1380
QY	1381	TGCCCTGACGAACCTCCCTTGAATTTAGCTTCTGTTTCATTTCCCTCAAAAAGATTTCCCA	1440
Db	1381	TGCCCTGACGAACCTCCCTTGAATTTAGCTTCTGTTTCATTTCCCTCAAAAAGATTTCCCA	1440

RESULT 3
US-09-503-444A-12
; Sequence 12, Application US/09503444A

Patent No. 6228594
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d1
US-09-503-444A-12

Query Match 100.0%; Score 1440; DB 4; Length 1440;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACCCTAGTGTGTTTCAACAAGCAGGTACCTTCTGCTGTAGAGAGAGAGA 60
DB 1 GGGGACACTGGATCACCCTAGTGTGTTTCAACAAGCAGGTACCTTCTGCTGTAGAGAGAGAGA 60
QY 61 ACTAAGTTCTGAAGACCTGTTGCTTTTACACAGGAAGTTTACTGGGCATCTCCTGAG 120
DB 61 ACTAAGTTCTGAAGACCTGTTGCTTTTACACAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CCTAGGCAATAGCTAGGCTGACTTCTGAGCCATCCCGCTTTCCCGCCGCCCAAAAG 180
DB 121 CCTAGGCAATAGCTAGGCTGACTTCTGAGCCATCCCGCTTTCCCGCCGCCCAAAAG 180
QY 181 AAGCGGAGATTTAAGCGGAGCTGCGGCCAGAGCTGGGAAATGGGCCGAGCCAGGC 240
DB 181 AAGCGGAGATTTAAGCGGAGCTGCGGCCAGAGCTGGGAAATGGGCCGAGCCAGGC 240
QY 241 CGGCGCTTCTCCTGATGCTTTTGCAGACCGGGTCTCTCAGGGGCTTGGGTT 300
DB 241 CGGCGCTTCTCCTGATGCTTTTGCAGACCGGGTCTCTCAGGGGCTTGGGTT 300
QY 301 CACACTCTCTGCACCTACCTCTTTCATGGGTGCTCAGACGAGACCTTGGTCTTCTTGT 360
DB 301 CACACTCTCTGCACCTACCTCTTTCATGGGTGCTCAGACGAGACCTTGGTCTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACCTGATGACGAGCTGTTGCTGTTCTATGATGAGAGTGC 420
DB 361 TTGAAGCTTTGGGCTACCTGATGACGAGCTGTTGCTGTTCTATGATGAGAGTGC 420
QY 421 GTGTGAGCCCGGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGGCTGAGC 480
DB 421 GTGTGAGCCCGGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGGCTGAGC 480
QY 481 TGAGTCAGAGTCTGAAAGGGTGGGATCATGTTTCACTGTTGACTTCTGACTATTATGG 540
DB 481 TGAGTCAGAGTCTGAAAGGGTGGGATCATGTTTCACTGTTGACTTCTGACTATTATGG 540
QY 541 AAATCAACAACACAGCAAGGAGTCCACACCTGCAAGTATCTTGGGCTGTGAATGC 600
DB 541 AAATCAACAACACAGCAAGGAGTCCACACCTGCAAGTATCTTGGGCTGTGAATGC 600
QY 601 AAGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
DB 601 AAGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
QY 661 AATCTGCCCTGACACACTGGATTGGAGAGCAGCAACCCAGGGGCTTGGCCCAACAGC 720
DB 661 AATCTGCCCTGACACACTGGATTGGAGAGCAGCAACCCAGGGGCTTGGCCCAACAGC 720
QY 721 TGGAGTGGAAAGGCACAAGATTGGGCCAGGCAAGACAGGGCTTACCTGGAGAGGACT 780
DB 721 TGGAGTGGAAAGGCACAAGATTGGGCCAGGCAAGACAGGGCTTACCTGGAGAGGACT 780
QY 781 GCCCTGCACAGCTGCAGCAGTTCCTGAGCTGGGAGAGTGTGTTGGACCAACAAGTGC 840
DB 781 GCCCTGCACAGCTGCAGCAGTTCCTGAGCTGGGAGAGTGTGTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACATCATGTGACCTCTTCACTGACCACTTACGGTGTGGG 900
DB 841 CTCCTTTGGTGAAGGTGACATCATGTGACCTCTTCACTGACCACTTACGGTGTGGG 900
QY 901 CTTTGAACCTACTACCCCGCAGACATCACCATGAAGTGGGCTGAAAGGATGAAGCAATGG 960
DB 901 CTTTGAACCTACTACCCCGCAGACATCACCATGAAGTGGGCTGAAAGGATGAAGCAATGG 960
QY 961 ATGCCAAGGAGTTGGAACCTAAAGACGTATTTGCCCAATGGGGATGGGACCTTACAGGGCT 1020
DB 961 ATGCCAAGGAGTTGGAACCTAAAGACGTATTTGCCCAATGGGGATGGGACCTTACAGGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACAGGTGGAGCACC 1080
DB 1021 GGATAACCTTGGCTGTACCCCTGGGGAAGACAGAGATATACGTACAGGTGGAGCACC 1080
QY 1081 CAGGCTGGATCAGGCCCTTCATGTTGATCTGGGAGCCCTCACCGCTCTGGCACCCTAGTCA 1140

Db 1081 CAGGCGCTGATCAGCCCTCATGTGATCTGGAGCCCTCACCGCTCGCACCTAGTCA 1140
QY 1141 TTGGAGTCATCAGTGAATTCGCTGTTTTTGTCTCATCTTGTTCATTGGAATTTTGTTC 1200
Db 1141 TTGGAGTCATCAGTGAATTCGCTGTTTTTGTCTCATCTTGTTCATTGGAATTTTGTTC 1200
QY 1201 TATATTAAAGAGAGGCGAGGTTCAAGAGGAGCCATGGGGCACATACGCTTAGCTGAAC 1260
Db 1201 TATATTAAAGAGAGGCGAGGTTCAAGAGGAGCCATGGGGCACATACGCTTAGCTGAAC 1260
QY 1261 GTGAGTGACACGCGCTGCACACTCCTGTGGGAGGAGCCATGGGGCACATACGCTTAGCTGAAC 1320
Db 1261 GTGAGTGACACGCGCTGCACACTCCTGTGGGAGGAGCCATGGGGCACATACGCTTAGCTGAAC 1320
QY 1321 AGGAGTGCAATTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCAATTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCCCTGACCAACTCCTTGATTTAGCCTTCTCTGTTTCATTTCCCTCAAAAAGATTTCCCCA 1440
Db 1381 TGCCCTGACCAACTCCTTGATTTAGCCTTCTCTGTTTCATTTCCCTCAAAAAGATTTCCCCA 1440

RESULT 4
US-08-652-265-10
; Sequence 10, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele

; LOCATION: replace(1066, "a")
; OTHER INFORMATION: /phenotype- "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
US-08-652-265-10

Query Match 99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTCTAGGAGAGAGA 60
Db 1 GGGGACACTGGATCACCCTAGTGTTCACAAGCAGGTACCTTCTGCTCTAGGAGAGAGA 60
QY 61 ACTAAAGTTCTGAAAGACCTGTTCTGTTTTTCAACAGGAAGTTTTTACTTGGGATCTCCCTGAG 120
Db 61 ACTAAAGTTCTGAAAGACCTGTTCTGTTTTTCAACAGGAAGTTTTTACTTGGGATCTCCCTGAG 120
QY 121 CCTAGGCAATAGCTGTAGGGTGACTTCTTGAGGCCATCCCGCTTTCCCGCGCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGTAGGGTGACTTCTTGAGGCCATCCCGCTTTCCCGCGCCCAAAAG 180
QY 181 AAGCGAGATTAAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGGCCCGGAGCCAGGC 240
Db 181 AAGCGAGATTAAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGGCCCGGAGCCAGGC 240
QY 241 CGGCGCTTCTCCTCTGATGCTTTTGACAGACCGCGGTCTTCAGAGGGCGCTTGTGCGTT 300
Db 241 CGGCGCTTCTCCTCTGATGCTTTTGACAGACCGCGGTCTTCAGAGGGCGCTTGTGCGTT 300
QY 301 CACACTCTGCACTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGGTCTTTCCCTTGT 360
Db 301 CACACTCTGCACTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGGTCTTTCCCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGATGACCAAGCTGTTTCGTTCTATGATGATGAGAGTCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGATGACCAAGCTGTTTCGTTCTATGATGATGAGAGTCGCC 420
QY 421 GTGTGAGCCCGCAACTCCATGCGTGTTCAGTAGAATTTCAAGCCAGATGGCTGCAGC 480
Db 421 GTGTGAGCCCGCAACTCCATGCGTGTTCAGTAGAATTTCAAGCCAGATGGCTGCAGC 480
QY 481 TGACTCAGAGCTGAAAGGGTGGGATCACATGTTTCACTGTTGACTTCTGACATATTATGG 540
Db 481 TGACTCAGAGCTGAAAGGGTGGGATCACATGTTTCACTGTTGACTTCTGACATATTATGG 540
QY 541 AAAATCAACACACAGCAAGGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAATGC 600
Db 541 AAAATCAACACACAGCAAGGAGTCCACACCCCTGCAGGTCATCCTGGGCTGTGAATGC 600
QY 601 AAGAAGACAACAGTACCGAGGGTACTTGAAGTACGGGTATGATGGCAGGACACCTTG 660
Db 601 AAGAAGACAACAGTACCGAGGGTACTTGAAGTACGGGTATGATGGCAGGACACCTTG 660
QY 661 AATTCTCCCTGCACACTCGATTGGAGAGCAGAGAACCCAGGGCTGCGCCACCAAGC 720
Db 661 AATTCTCCCTGCACACTCGATTGGAGAGCAGAGAACCCAGGGCTGCGCCACCAAGC 720
QY 721 TGGAGTGGAAAGGCACAAGATTCGGGCCAGGAGCAGAACAGGGCTTACCTGAGAGGACT 780
Db 721 TGGAGTGGAAAGGCACAAGATTCGGGCCAGGAGCAGAACAGGGCTTACCTGAGAGGACT 780
QY 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGAGTGTGGACCAACAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGAGTGTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGAACACTCTACGCTGTGGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGAACACTCTACGCTGTGGG 900
QY 901 CCTTGAACCTACTACCCCAAGAACATCACCATGAAGTGGCTGAAGGATTAAGCACCCTAATGG 960
Db 901 CCTTGAACCTACTACCCCAAGAACATCACCATGAAGTGGCTGAAGGATTAAGCACCCTAATGG 960

Db 901 CCTTGAACCTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGGATGAAGCAGCAATGG 960
Qy 961 ATGCCAAGAGTTCGAACCTAAGAGGTATTGCCAAATGGGGATGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGAGTTCGAACCTAAGAGGTATTGCCAAATGGGGATGGACCTACCAAGGCT 1020
Qy 1021 GGATAACCTTGGCTGTACCCCCCTGGGGAAGACAGAGATATACGTACCAGGTGGAGCACC 1080
Db 1021 GGATAACCTTGGCTGTACCCCCCTGGGGAAGACAGAGATATACGTACCAGGTGGAGCACC 1080
Qy 1081 CAGGCCCTGGATCAGCCCTCATATTGATCTGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db 1081 CAGGCCCTGGATCAGCCCTCATATTGATCTGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Qy 1141 TTGGAGTCATCAGTGAATTCGTTTTTGTGCTCATCTTGTTCATTGSAATTTTGTCA 1200
Db 1141 TTGGAGTCATCAGTGAATTCGTTTTTGTGCTCATCTTGTTCATTGSAATTTTGTCA 1200
Qy 1201 TAATATTAAGGAAGGCGAGGGTTCAAGAGAGGCCATGGGGCACTACGTTTAGCTGAAC 1260
Db 1201 TAATATTAAGGAAGGCGAGGGTTCAAGAGAGGCCATGGGGCACTACGTTTAGCTGAAC 1260
Qy 1261 GTGAGTGACACGCGCTGCACACTCAGTCTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTGACACGCGCTGCACACTCAGTCTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Qy 1321 AGGAGCTGATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGCTGATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Qy 1381 TGCCCTGACGAACCTCTTGATTTTAGCCTTCTGTTCATTTCCCTCAAAAAGATTTCCCA 1440
Db 1381 TGCCCTGACGAACCTCTTGATTTTAGCCTTCTGTTCATTTCCCTCAAAAAGATTTCCCA 1440

RESULT 5

US-08-652-265-11
; Sequence 11, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(408, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION: /label= 24d2
; US-08-652-265-11

Query Match 99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTAGGAGAGAGA 60
Db 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTAGGAGAGAGA 60
Qy 61 ACTAAAGTCTGAAAGACCTGTTGCTTTTCCACCAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTCTGAAAGACCTGTTGCTTTTCCACCAGGAAGTTTACTGGGCATCTCCTGAG 120
Qy 121 CCTAGGCAATAGCTGTAGGTGACTTCTTGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGTAGGTGACTTCTTGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
Qy 181 AAGCGGAGATTTAAGCGGAGCTGCGGCCAGAGCTGGGAAATGGGCCCGCGAGCCAGGC 240
Db 181 AAGCGGAGATTTAAGCGGAGCTGCGGCCAGAGCTGGGAAATGGGCCCGCGAGCCAGGC 240
Qy 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGCTTCAGAGGCGGCTTGCCTGCTT 300
Db 241 CGGCGCTTCTCCTCTGATGCTTTTGCAGACCGCGGCTTCAGAGGCGGCTTGCCTGCTT 300
Qy 301 CACACTCTCTGCACTACCTCTTCATGGTGCCTCAGACGAGGACCTTGGTCTTCTCCTTGT 360
Db 301 CACACTCTCTGCACTACCTCTTCATGGTGCCTCAGACGAGGACCTTGGTCTTCTCCTTGT 360
Qy 361 TTGAAGCTTTGGGCTAGCTGGATGACCAAGCTGTTTCGTTCTTATGATGATGAGAGTCCGC 420
Db 361 TTGAAGCTTTGGGCTAGCTGGATGACCAAGCTGTTTCGTTCTTATGATGATGAGAGTCCGC 420
Qy 421 GTGTGGGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Qy 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAAGGGTGGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
Qy 541 AAAATCAACAACACAGCAAGGAGTCCACACCCCTGCAGGTCACTCCTGGGTGTGAAATGC 600
Db 541 AAAATCAACAACACAGCAAGGAGTCCACACCCCTGCAGGTCACTCCTGGGTGTGAAATGC 600
Qy 601 AAGAAACACAGTACCGAGGGCTACTTGGAAAGTACGGGTATGATGGCAGGACCACCTTG 660
Db 601 AAGAAACACAGTACCGAGGGCTACTTGGAAAGTACGGGTATGATGGCAGGACCACCTTG 660
Qy 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAACCCAGGCTTGGCCCAACCAAGC 720
Db 661 AATTCTGCCCTGACACACTGGATTGGAGAGCAGCAACCCAGGCTTGGCCCAACCAAGC 720
Qy 721 TGGAGTGGGAAAGGCACAAGATTCCGGCCAGGAGCAACAGGCGCTTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAAGGCACAAGATTCCGGCCAGGAGCAACAGGCGCTTACCTGGAGAGGACT 780

```
Qy 781 GCCCTGCACACTGCAGCAGTTCCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 GCCCTGCACACTGCAGCAGTTCCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGTGC 840
Qy 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAAGTACCACTCTACGGTGTGCGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAAGTACCACTCTACGGTGTGCGG 900
Qy 901 CTTGAACCTACTACCCAGAACATCAACATGAAGTGGCTGAAGATAAGCAGCAATGG 960
Db 901 CTTGAACCTACTACCCAGAACATCAACATGAAGTGGCTGAAGATAAGCAGCAATGG 960
Qy 961 ATGCCAAGAGGTTCGAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACACAGGCT 1020
Db 961 ATGCCAAGAGGTTCGAACCTAAAGACGTATTGCCCAATGGGATGGGACCTACACAGGCT 1020
Qy 1021 GGATAACCTTGGCTGTACCCCTCGGGGAAGAGCAGAGATATACGTACAGGTGGAGCAC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTCGGGGAAGAGCAGAGATATACGTACAGGTGGAGCAC 1080
Qy 1081 CAGGCTGGATCAGCCCTCATTTGATCTGGAGCCCTCACCGTCTGGCACCCCTAGTCA 1140
Db 1081 CAGGCTGGATCAGCCCTCATTTGATCTGGAGCCCTCACCGTCTGGCACCCCTAGTCA 1140
Qy 1141 TTGGAGTCATCAGTGAATTCGTGTTTGTGCTCATCTTCTATTGGAATTTGTTCA 1200
Db 1141 TTGGAGTCATCAGTGAATTCGTGTTTGTGCTCATCTTCTATTGGAATTTGTTCA 1200
Qy 1201 TAATATTAAAGGAAGGAGGAGGTTCAAGAGGAGGACCATGGGACCTACCTTTAGTGAAC 1260
Db 1201 TAATATTAAAGGAAGGAGGAGGTTCAAGAGGAGGACCATGGGACCTACCTTTAGTGAAC 1260
Qy 1261 GTGAGTGACACAGCCTGCAGACTCATCTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
Db 1261 GTGAGTGACACAGCCTGCAGACTCATCTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
Qy 1321 AGGGAGTCATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGGAGTCATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Qy 1381 TGCTGACGAACCTCTGATTTAGCTTCTCTGTTTCAATTTCTCAAAAAGATTTCCTCCA 1440
Db 1381 TGCTGACGAACCTCTGATTTAGCTTCTCTGTTTCAATTTCTCAAAAAGATTTCCTCCA 1440
```

```
RESULT 6
US-08-834-497A-10
; Sequence 10, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gairke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Polissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1066, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION: /label= 24d1
; US-08-834-497A-10
```

```
Query Match 99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTAACCTTCTCTCTAGAGAGAGAGA 60
Db 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTAACCTTCTCTCTAGAGAGAGAGA 60
Qy 61 ACTAAGTTCCTGAAGACCTGTTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAGTTCCTGAAGACCTGTTGCTTTTCCACAGGAAGTTTACTGGGCATCTCCTGAG 120
Qy 121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGTTTCCCGCCCCCAAAAG 180
Qy 181 AAGCGGAGATTAAACGGGGACGTGCGCCAGAGCTGGGAAATGGCGCGGAGCCAGGC 240
Db 181 AAGCGGAGATTAAACGGGGACGTGCGCCAGAGCTGGGAAATGGCGCGGAGCCAGGC 240
Qy 241 CGGGCTTCTCCTCTGATGCTTTTGACAGCCGCGGTCTCAGGGGCGCTTGTGCGGTT 300
Db 241 CGGGCTTCTCCTCTGATGCTTTTGACAGCCGCGGTCTCAGGGGCGCTTGTGCGGTT 300
Qy 301 CACACTCTCTGACTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Db 301 CACACTCTCTGACTACCTCTTCATGGGTGCTCAGAGCAGGACCTTGGTCTTTCCTTGT 360
Qy 361 TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTGCGTCTTATGATGAGAGAGTGC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTGCGTCTTATGATGATGAGAGTGC 420
```

```
Oy 421 GTGTGGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Oy 481 TGAGTCAGATCTGAAGAGGTGGGATACATGTTCACTGTTGACTTCTGGACTATATATGG 540
Db 481 TGAGTCAGATCTGAAGAGGTGGGATACATGTTCACTGTTGACTTCTGGACTATATATGG 540
Oy 541 AAAATCAACACACAGCAAGAGTCCACACCCCTGCAGCTCATCTGGGCTGTGAATGC 600
Db 541 AAAATCAACACACAGCAAGAGTCCACACCCCTGCAGCTCATCTGGGCTGTGAATGC 600
Oy 601 AAGAGACACAGTACCGAGGGTACTGGAGTACGGGTATGATGGGAGGACCACTTGG 660
Db 601 AAGAGACACAGTACCGAGGGTACTGGAGTACGGGTATGATGGGAGGACCACTTGG 660
Oy 661 AATTCTGCCCTGCACACATGATTGGAGAGCAGCAACCCAGGGCCTGGCCCAACAGC 720
Db 661 AATTCTGCCCTGCACACATGATTGGAGAGCAGCAACCCAGGGCCTGGCCCAACAGC 720
Oy 721 TGGAGTGGGAAGGCACACAGATTTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Db 721 TGGAGTGGGAAGGCACACAGATTTCGGGCCAGGCAGAACAGGGCCTACCTGGAGAGGGACT 780
Oy 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCACCAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGCAGTTGCTGGAGCTGGGAGAGGTGTTTGGACCACCAAGTGC 840
Oy 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTTACGGGTGTCGGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTTACGGGTGTCGGG 900
Oy 901 CTTTGAACCTACTACCCCAAGCATCACCATGAAGTGGCTGAAGGATAAGCAAGCAATGG 960
Db 901 CTTTGAACCTACTACCCCAAGCATCACCATGAAGTGGCTGAAGGATAAGCAAGCAATGG 960
Oy 961 ATGCCAAGGAGTTCGAACCTTAAGAGATATGCCCCAATGGGATGGGACCTACCAAGGCT 1020
Db 961 ATGCCAAGGAGTTCGAACCTTAAGAGATATGCCCCAATGGGATGGGACCTACCAAGGCT 1020
Oy 1021 GGATAACCTTGGCTGTATCCCTCCCTGGGGAAGCAGACAGATATACGTACCAAGTGGAGCAC 1080
Db 1021 GGATAACCTTGGCTGTATCCCTCCCTGGGGAAGCAGACAGATATACGTACCAAGTGGAGCAC 1080
Oy 1081 CAGCCCTGGATCAGCCCTCATTTGATCTGGAGAGCCCTCACGGTCTGGCACCCTAGTCA 1140
Db 1081 CAGCCCTGGATCAGCCCTCATTTGATCTGGAGAGCCCTCACGGTCTGGCACCCTAGTCA 1140
Oy 1141 TTGAGTCACTCAGTGAATTCGTTTGTCTGCTCATCTGTTTCATTGGAATTTGTTCA 1200
Db 1141 TTGAGTCACTCAGTGAATTCGTTTGTCTGCTCATCTGTTTCATTGGAATTTGTTCA 1200
Oy 1201 TAATATTAAAGAGAGGACGGTTCAAGAGAGGCCATGGGGCACTACGTCCTTAGCTGAAC 1260
Db 1201 TAATATTAAAGAGAGGACGGTTCAAGAGAGGCCATGGGGCACTACGTCCTTAGCTGAAC 1260
Oy 1261 GTGAGTGACAGCGCCTGCAGACTCAGTCTGGAGAGGACAAACACTAGAGACTCAAG 1320
Db 1261 GTGAGTGACAGCGCCTGCAGACTCAGTCTGGAGAGGACAAACACTAGAGACTCAAG 1320
Oy 1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGGAGTGCATTTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Oy 1381 TGCCTGAGCAACTCCTTGATTTTAGCCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCA 1440
Db 1381 TGCCTGAGCAACTCCTTGATTTTAGCCTTCTCTGTTTCATTTCTCAAAAAGATTTCCCA 1440
```

RESULT 7

US-08-834-497A-11

; Sequence 11, Application US/08834497A

```
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/632,673
; FILING DATE: 16-APR-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(408, "g")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d2
; US-08-834-497A-11
```

Query Match 99.9%; Score 1438.4; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
Db 1 GGGGACACTGGATCACCCTAGTGTTCACAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60

QY 61 ACTAAGTTCTGAAGACCTGTTGGCTTTTTCACCAAGAAAGTTTACTGGGCATCTCCTGAG 120
|||||
Db 61 ACTAAGTTCTGAAGACCTGTTGGCTTTTTCACCAAGAAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CCTAGCAATAGCTGTAGGGTCACTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
|||||
Db 121 CCTAGCAATAGCTGTAGGGTCACTTCTGGAGCCATCCCGTTTCCCGCCGCCCAAAAG 180
QY 181 AAGCGGAGATTAAAGGGGACCTGGGGCAGAGCTGGGAAATGGGCCCGCAGCAGGC 240
|||||
Db 181 AAGCGGAGATTAAAGGGGACCTGGGGCAGAGCTGGGAAATGGGCCCGCAGCAGGC 240
QY 241 CGGCGCTTCTCCTGATGCTTTTTCAGACCGCGGTCTCGAGGGGCGCTTGTGCGGTT 300
|||||
Db 241 CGGCGCTTCTCCTGATGCTTTTTCAGACCGCGGTCTCGAGGGGCGCTTGTGCGGTT 300
QY 301 CACACTCTCTGACACCTCTCATGGGTGCTCAGACGAGCACTTGGCTTTCTCTGTT 360
|||||
Db 301 CACACTCTCTGACACCTCTCATGGGTGCTCAGACGAGCACTTGGCTTTCTCTGTT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTGCTGTTCTATGATGATGAGAGTCGC 420
|||||
Db 361 TTGAAGCTTTGGGCTACGTGGATGACCACTGTTGCTGTTCTATGATGATGAGAGTCGC 420
QY 421 GTGTGAGCCCCGAACCTCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
|||||
Db 421 GTGTGAGCCCCGAACCTCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TCAGTCAGAGTCTGAAGGGTGGGATCAGATGTTCACTGTTGACTTCTGGACTATATGG 540
|||||
Db 481 TCAGTCAGAGTCTGAAGGGTGGGATCAGATGTTCACTGTTGACTTCTGGACTATATGG 540
QY 541 AAAATCACAAACACAGCAGAGTCCACACCTCGAGGTCTCTGAGGCTGTGAATGC 600
|||||
Db 541 AAAATCACAAACACAGCAGAGTCCACACCTCGAGGTCTCTGAGGCTGTGAATGC 600
QY 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG 660
|||||
Db 601 AAGAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG 660
QY 661 AATTCTGCCCTCACACACTGGAATGGAGAGCAGCAACCCAGGGCTGSCCCCAACAAGC 720
|||||
Db 661 AATTCTGCCCTCACACACTGGAATGGAGAGCAGCAACCCAGGGCTGSCCCCAACAAGC 720
QY 721 TGAAGTGGAAAGGCACAAGATTTCGGCCAGGCAGCAACAGGGCTACCTGGAGAGGACT 780
|||||
Db 721 TGAAGTGGAAAGGCACAAGATTTCGGCCAGGCAGCAACAGGGCTACCTGGAGAGGACT 780
QY 781 GCCCTGCACAGCTGCAGCAGTTCGTGGAGCTGGGAGAGGTGTTTGGACCACAAGTGC 840
|||||
Db 781 GCCCTGCACAGCTGCAGCAGTTCGTGGAGCTGGGAGAGGTGTTTGGACCACAAGTGC 840
QY 841 CTCCTTTGGTGAAGTGACACATCATGTGACCTTTCAGTGACCACTACCGTGTCCGG 900
|||||
Db 841 CTCCTTTGGTGAAGTGACACATCATGTGACCTTTCAGTGACCACTACCGTGTCCGG 900
QY 901 CCTTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
|||||
Db 901 CCTTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGGATAAGCAGCCAATGG 960
QY 961 ATGCCAAGAGTTCGAACCTTAAGAGGTATTGCCAATGGGATGGGACCTACCGAGGGT 1020
|||||
Db 961 ATGCCAAGAGTTCGAACCTTAAGAGGTATTGCCAATGGGATGGGACCTACCGAGGGT 1020
QY 1021 GGATACCTTGGCTGTACCCCTGGGAGAGCAGAGATATAGTACAGGTGGAGCACC 1080
|||||
Db 1021 GGATACCTTGGCTGTACCCCTGGGAGAGCAGAGATATAGTACAGGTGGAGCACC 1080
QY 1081 CAGGCTGTGATCAGCCCTCATTTGATCTGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
|||||
Db 1081 CAGGCTGTGATCAGCCCTCATTTGATCTGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140

QY 1141 TTGGAGTCATCAGTGAATTGCTGTTTTTGTGCTCATCTTGTTCATTGGAATTTCTTCA 1200
|||||
Db 1141 TTGGAGTCATCAGTGAATTGCTGTTTTTGTGCTCATCTTGTTCATTGGAATTTCTTCA 1200
QY 1201 TAATATTAAGGAAGAGCGAGGTTCAAGAGGAGCCATGGGCACCTACGTTCTTAGCTGAAC 1260
|||||
Db 1201 TAATATTAAGGAAGAGCGAGGTTCAAGAGGAGCCATGGGCACCTACGTTCTTAGCTGAAC 1260
QY 1261 GTGAGTGCACACGCGCTGCAGACTCCTGTGGGAAGGAGACAAAACCTAGAGACTCAAAAG 1320
|||||
Db 1261 GTGAGTGCACACGCGCTGCAGACTCCTGTGGGAAGGAGACAAAACCTAGAGACTCAAAAG 1320
QY 1321 AGGGAGTCATTTATCAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
|||||
Db 1321 AGGGAGTCATTTATCAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCCTGACCAACTCCTTGATTTTAGCTTCTCTGTGTTTCATTTCTCAAAAAGATTTCCCCA 1440
|||||
Db 1381 TGCCTGACCAACTCCTTGATTTTAGCTTCTCTGTGTTTCATTTCTCAAAAAGATTTCCCCA 1440
RESULT 8
US-09-503-444A-10
; Sequence 10, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gairke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/652,265
; FILING DATE: 23-May-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,673
; FILING DATE: 16-Apr-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,912
; FILING DATE: 04-Apr-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0088-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222...1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d1
US-09-503-444A-10

Query Match 99.9%; Score 1438.4; DB 4; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGTAAGTCTGCTGTAGGAGAGAGAGA 60
Db 1 GGGGACACTGGATCAGCTAGTGTTCACAGCAGGTAAGTCTGCTGTAGGAGAGAGAGA 60
QY 61 ACTAAAGTTCTGAAGACCTGTGCTTTTTCACAGGAAGTTTACTGGGCATCTCCTGAG 120
Db 61 ACTAAAGTTCTGAAGACCTGTGCTTTTTCACAGGAAGTTTACTGGGCATCTCCTGAG 120
QY 121 CCTAGGCAATAGCTGTAGGTAAGTCTGAGGCAATCCCGCTTTCCCGCCCGCCCAAG 180
Db 121 CCTAGGCAATAGCTGTAGGTAAGTCTGAGGCAATCCCGCTTTCCCGCCCGCCCAAG 180
QY 181 AAGCGGAGATTTAAGCGGACGTGGGCGCAGAGCTGGGGAATGGGCGCGGACGAGGC 240
Db 181 AAGCGGAGATTTAAGCGGACGTGGGCGCAGAGCTGGGGAATGGGCGCGGACGAGGC 240
QY 241 CGCGCGTCTCTCTCTGTATGCTTTTTCAGACCGCGGTCTGCGAGGGCGCTTGCTGCTT 300
Db 241 CGCGCGTCTCTCTCTGTATGCTTTTTCAGACCGCGGTCTGCGAGGGCGCTTGCTGCTT 300
QY 301 CACACTCTGTGCACTACCTCTTCATGGTGTGCTTCAGACGAGACCTTGCTTCTTCTGT 360
Db 301 CACACTCTGTGCACTACCTCTTCATGGTGTGCTTCAGACGAGACCTTGCTTCTTCTGT 360
QY 361 TTGAGAGCTTTGGGCTACGTGGATGACGAGCTGTGCTGTTCTATGATGATGAGTGGCC 420
Db 361 TTGAGAGCTTTGGGCTACGTGGATGACGAGCTGTGCTGTTCTATGATGATGAGTGGCC 420
QY 421 GTGTGGAGCCCGCACTCCATGGTTCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTGTGGAGCCCGCACTCCATGGTTCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCAGATGTTTCACTGTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCAGATGTTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 AAAATCACAACACAGCAGAGGTCCACACCTGCAAGTCACTGCGGCTGTGAATGC 600
Db 541 AAAATCACAACACAGCAGAGGTCCACACCTGCAAGTCACTGCGGCTGTGAATGC 600
QY 601 AAGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
Db 601 AAGAAGACAACAGTACCGAGGCTACTGGAAGTACGGGTATGATGGGAGGACCACTTG 660
QY 661 AATTCTGCCCTGACACTGATTTGGAGAGCAGACCAAGGCGCTGGCCCAAGC 720
Db 661 AATTCTGCCCTGACACTGATTTGGAGAGCAGACCAAGGCGCTGGCCCAAGC 720
QY 721 TGGAGTGGGAAGGCACAAGATTGGGCGCAGCAGAGGCGCTTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAGGCACAAGATTGGGCGCAGCAGAGGCGCTTACCTGGAGAGGACT 780
QY 781 GCGCTGCAACAGTGCAGCAGTGTCTGAGGCTGGGAGAGGTGTTTGGACCAACAGTGC 840
Db 781 GCGCTGCAACAGTGCAGCAGTGTCTGAGGCTGGGAGAGGTGTTTGGACCAACAGTGC 840

QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGGG 900
Db 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTTCAGTGACCACTCTACGGTGTGGG 900
QY 901 CTTTGAACACTACTACCCCGGAGAACATCAACATGAAGTGGCTGAAGGATGAAGCAGCAATGG 960
Db 901 CTTTGAACACTACTACCCCGGAGAACATCAACATGAAGTGGCTGAAGGATGAAGCAGCAATGG 960
QY 961 ATGCCAAGGAGTTGGAACCTTAAAGACGTATTGCCCAATGGGATGGGATGGGATGGGCT 1020
Db 961 ATGCCAAGGAGTTGGAACCTTAAAGACGTATTGCCCAATGGGATGGGATGGGATGGGCT 1020
QY 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGAGCAGAGATATACGTACCAAGGTGGAGCACC 1080
Db 1021 GGATAACCTTTGGCTGTACCCCTGGGGAAGAGCAGAGATATACGTACCAAGGTGGAGCACC 1080
QY 1081 CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
Db 1081 CAGCCCTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
QY 1141 TTGGAGTCATCAGTGAATGCTGTTTTTCTGCTCATCTTTGTTTCATTGGAATTTTGTTC 1200
Db 1141 TTGGAGTCATCAGTGAATGCTGTTTTTCTGCTCATCTTTGTTTCATTGGAATTTTGTTC 1200
QY 1201 TAATATTAAAGAAAGAGAGCGGTTCAAGAGGAGCCATGGGCGCACTACGTTCTTACGTTGAAC 1260
Db 1201 TAATATTAAAGAAAGAGAGCGGTTCAAGAGGAGCCATGGGCGCACTACGTTCTTACGTTGAAC 1260
QY 1261 GTGAGTCACAGCAGCCCTGCAGACTCACTGTGGGAAGGAGACAAAAGTAGAGACTCAAG 1320
Db 1261 GTGAGTCACAGCAGCCCTGCAGACTCACTGTGGGAAGGAGACAAAAGTAGAGACTCAAG 1320
QY 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCGTACGAACTCTTGATTTTAGCCTTCTCTGTTTCATTTCCTCAAAAGATTTCCCA 1440
Db 1381 TGCGTACGAACTCTTGATTTTAGCCTTCTCTGTTTCATTTCCTCAAAAGATTTCCCA 1440

RESULT 9
US-09-503-444A-11
; Sequence 11, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d2
US-09-503-444A-11

Query Match 99.9%; Score 1438.4; DB 4; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1439; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	GGGGACACTGGATACACCTAGTGTTCACAGCAGGTACTCTGTCTGTAGGAGAGAGAGA	60
Db	1	GGGGACACTGGATACACCTAGTGTTCACAGCAGGTACTCTGTCTGTAGGAGAGAGAGA	60
QY	61	ACTAAGTTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCACTCCTGAG	120
Db	61	ACTAAGTTCTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCACTCCTGAG	120
QY	121	CCTAGGCAATAGCTGTAGGGTACTCTTGGAGCCATCCCGCTTTCCCGCCGCCCAAAAG	180
Db	121	CCTAGGCAATAGCTGTAGGGTACTCTTGGAGCCATCCCGCTTTCCCGCCGCCCAAAAG	180
QY	181	AAGCGGAGATTTAAGGGGACGTGGCCAGAGCTGGGGAATGGGCCCGGAGCCAGGC	240
Db	181	AAGCGGAGATTTAAGGGGACGTGGCCAGAGCTGGGGAATGGGGCAATGGGGCGGAGCCAGGC	240
QY	241	CGGCGCTTCTCTCTGTATGCTTTTGCAGACCGCGTCTTGCAGGGCGCTTGGTGCCTT	300
Db	241	CGGCGCTTCTCTCTGTATGCTTTTGCAGACCGCGGTCTTGCAGGGCGCTTGGTGCCTT	300
QY	301	CACACTCTGTGACACTCTTCTATGGGTGCTTCAGAGCAGGACCTTGGTCTTCTCTGT	360
Db	301	CACACTCTGTGACACTCTTCTATGGGTGCTTCATGGGTGCTTCAGAGCAGGACCTTGGTCTTCTGT	360
QY	361	TTGAGCTTTGGGCTACGTGGATACCAAGCTGTTCGTGTTTATGATGATGAGAGTCGCC	420
Db	361	TTGAGCTTTGGGCTACGTGGATACCAAGCTGTTCGTGTTTATGATGATGAGAGTCGCC	420
QY	421	GTCTGGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC	480
Db	421	GTCTGGAGCCCCGAACCTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC	480
QY	481	TGAGTTCAGAGTCTGAAGGGTGGGATCACATGTTCCACTGTGTGACTTCTGGACTATTATGG	540

Db	481	TGAGTTCAGAGTCTGAAGGGTGGGATCACATGTTTCACTGTTGACTTCTGGACTATTATGG	540
QY	541	AAATATCACACCAACAGAGTCCACACACCTGCGAGTCACTCTGGGCTGTCAATATGC	600
Db	541	AAATATCACACCAACAGAGTCCACACACCTGCGAGTCACTCTGGGCTGTCAATATGC	600
QY	601	AAGAAGACAAACAGTACCGAGGGCTACTTGGAAAGTACGGGTATGATGGCAGGACCACTTG	660
Db	601	AAGAAGACAAACAGTACCGAGGGCTACTTGGAAAGTACGGGTATGATGGCAGGACCACTTG	660
QY	661	AATTCTGCCCTGCACACTGGATTGGAGAGCAGCAACCCAGGGCTGCCCCACCAAGC	720
Db	661	AATTCTGCCCTGCACACTGGATTGGAGAGCAGCAACCCAGGGCTGCCCCACCAAGC	720
QY	721	TGGAGTGGGAAAGGCACAAGATTGCGGCCAGGAGAACAGGGCTACTCTGAGAGGAGCT	780
Db	721	TGGAGTGGGAAAGGCACAAGATTGCGGCCAGGAGAACAGGGCTACTCTGAGAGGAGCT	780
QY	781	GCCCTGCACAGCTGCACAGTGTGCTGGAGCTGGGGAGAGTGTTTGGACCAACAAGTGC	840
Db	781	GCCCTGCACAGCTGCACAGTGTGCTGGAGCTGGGGAGAGTGTTTGGACCAACAAGTGC	840
QY	841	CTCCTTTGGTGAAGTGCACATCATCTGACCTCTTCCAGTGACCCTCTACGGTGTGCGG	900
Db	841	CTCCTTTGGTGAAGTGCACATCATCTGACCTCTTCCAGTGACCCTCTACGGTGTGCGG	900
QY	901	CCTTGAACCTACTACCCCGCAGAACATCACCATGAAGTGGCTGAAGATAAGCAGCAATGG	960
Db	901	CCTTGAACCTACTACCCCGCAGAACATCACCATGAAGTGGCTGAAGATAAGCAGCAATGG	960
QY	961	ATGCCAAGGAGTTGCAACCTTAAGACGTATTGCCAATGGGGATGGACCTACAGGGCT	1020
Db	961	ATGCCAAGGAGTTGCAACCTTAAGACGTATTGCCAATGGGGATGGACCTACAGGGCT	1020
QY	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGCAGAGATATACGTACACAGTGGAGCACC	1080
Db	1021	GGATAACCTTGGCTGTACCCCTGGGGAAGCAGAGATATACGTACACAGTGGAGCACC	1080
QY	1081	CAGCCCTGGATCAGCCCTCATTTGTATCTGGGAGCCCTCACCCTCTGGCACCCTAGTCA	1140
Db	1081	CAGCCCTGGATCAGCCCTCATTTGTATCTGGGAGCCCTCACCCTCTGGCACCCTAGTCA	1140
QY	1141	TTGAGTTCATCAGTGGAAATTTGCTGTTTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA	1200
Db	1141	TTGAGTTCATCAGTGGAAATTTGCTGTTTTTTCGTCATCTTGTTCATTTGGAATTTTGTCA	1200
QY	1201	TAATATTAAAGGAAGAGCAGGGTTCAAGAGAGCCATGGGGCCTACGTCTTAGCTGAAC	1260
Db	1201	TAATATTAAAGGAAGAGCAGGGTTCAAGAGAGCCATGGGGCCTACGTCTTAGCTGAAC	1260
QY	1261	GTGAGTGACACGAGCCTGCAGACTCACTGTGGGAAGAGACAAACTAGAGACTCAAG	1320
Db	1261	GTGAGTGACACGAGCCTGCAGACTCACTGTGGGAAGAGACAAACTAGAGACTCAAG	1320
QY	1321	AGGAGTGCATTTATGAGTCTTCATGTTTCAGSAGAGTGTGAACCTAAACATAGAAAT	1380
Db	1321	AGGAGTGCATTTATGAGTCTTCATGTTTCAGSAGAGTGTGAACCTAAACATAGAAAT	1380
QY	1381	TGCTGACGAACCTCTTGATTTTAGCCTTCTGTTTCATTTCCCTCAAAAAGATTTCCCCA	1440
Db	1381	TGCTGACGAACCTCTTGATTTTAGCCTTCTGTTTCATTTCCCTCAAAAAGATTTCCCCA	1440

RESULT 10
US-08-652-265-9
; Sequence 9, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David

APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,265
FILING DATE: 23-MAY-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 17957-000500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1440 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 222..1268
FEATURE:
NAME/KEY: allele
LOCATION: replace(408, "c")
OTHER INFORMATION: (phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(414, "a")
OTHER INFORMATION: (phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d7
FEATURE:
NAME/KEY: allele
LOCATION: replace(1066, "g")
OTHER INFORMATION: (phenotype= "normal or wild-type"
OTHER INFORMATION: (unaffected)
OTHER INFORMATION: /label= 24d1

US-08-652-265-9

Query Match 99.8%; Score 1436.8; DB 3; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGACACTGGATACCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGGAGAGAGA 60
DB 1 GGGACACTGGATACCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGGAGAGAGA 60
QY 61 ACTAAGTTCTGAAGACCTGTGCTTTTCCACGAGAGTTTACTGGGCATCTCCTGAG 120
DB 61 ACTAAGTTCTGAAGACCTGTGCTTTTCCACGAGAGTTTACTGGGCATCTCCTGAG 120
QY -121 CCTAGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGCTTCCCGCCCCCAAAAG 180
DB 121 CCTAGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCGCTTCCCGCCCCCAAAAG 180

QY 181 AAGCGAGATTTAAGCGGGACGTCGGCCAGAGCTGGGAAATGGCCCGCGAGCCAGGC 240
DB 181 AAGCGAGATTTAAGCGGGACGTCGGCCAGAGCTGGGAAATGGCCCGCGAGCCAGGC 240
QY 241 CGGGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTCGAGGGCGCTTCTCGGTT 300
DB 241 CGGGCTTCTCCTCTGATGCTTTTGCAGACCGCGGTCTCGAGGGCGCTTCTCGGTT 300
QY 301 CACACTCTGCACTACCTCTTCATGGTCCCTCAGACAGGACCTTGGTCTTCTTGT 360
DB 301 CACACTCTGCACTACCTCTTCATGGTCCCTCAGACAGGACCTTGGTCTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTCGTGTCTTATGATGAGAGTCGCC 420
DB 361 TTGAAGCTTTGGGCTACGTGGATGACCAAGCTGTTCGTGTCTTATGATGAGAGTCGCC 420
QY 421 GTGTGGAGCCCCGAATCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTCGAGC 480
DB 421 GTGTGGAGCCCCGAATCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTCGAGC 480
QY 481 TGAGTCAGAGCTCTCAAGGGTGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
DB 481 TGAGTCAGAGCTCTCAAGGGTGGATCACATGTTCACTGTTGACTTCTGGACTATTATGG 540
QY 541 -AAATCACAACCACAGCAAGGAGTCCACACCTCGAGGTCTCCTGGGCTGTGAATGC 600
DB 541 -AAATCACAACCACAGCAAGGAGTCCACACCTCGAGGTCTCCTGGGCTGTGAATGC 600
QY 601 AGAAGACACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG 660
DB 601 AGAAGACACAGTACCGAGGGCTACTGGAAGTACGGGTATGATGGCAGGACCACTTG 660
QY 661 AATTCTGCCCTGCACACTCTGATTTGGAGACAGCAACCCAGGGCTTGCCCCCAACAGC 720
DB 661 AATTCTGCCCTGCACACTCTGATTTGGAGACAGCAACCCAGGGCTTGCCCCCAACAGC 720
QY 721 TGGAGTGGGAAAGGCACAAGATTGGGCCAGGAGAGAACAGGGCTTACCTGGAGAGGACT 780
DB 721 TGGAGTGGGAAAGGCACAAGATTGGGCCAGGAGAGAACAGGGCTTACCTGGAGAGGACT 780
QY 781 GCCTGTGACAGCTGCAGCAGTCTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
DB 781 GCCTGTGACAGCTGCAGCAGTCTGCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
DB 841 CTCCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGG 900
QY 901 CCTTGAACCTACTACCCCGCAACATCACCATGAAGTGGCTGAAGGATAGACAGCCAATGG 960
DB 901 CCTTGAACCTACTACCCCGCAACATCACCATGAAGTGGCTGAAGGATAGACAGCCAATGG 960
QY 961 ATGCCAAGGAGTTCGAACCTTAAGACGTATTGCCCAATGGGATGGAGCTTACCAAGGCT 1020
DB 961 ATGCCAAGGAGTTCGAACCTTAAGACGTATTGCCCAATGGGATGGAGCTTACCAAGGCT 1020
QY 1021 GGATAACCTTTGGTGTATACCCCTGGGGAAGACAGAGATATACGTACCAAGGTGGAGCAC 1080
DB 1021 GGATAACCTTTGGTGTATACCCCTGGGGAAGACAGAGATATACGTACCAAGGTGGAGCAC 1080
QY 1081 CAGGCTTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
DB 1081 CAGGCTTGGATCAGCCCTCATTTGTGATCTGGGAGCCCTCACCGTCTGGCACCCTAGTCA 1140
QY 1141 TTGAGTTCATCAGTGGAAATTCGTTTTTCTGTCATCTTTGTTTCAATGGAAATTTGTTCA 1200
DB 1141 TTGAGTTCATCAGTGGAAATTCGTTTTTCTGTCATCTTTGTTTCAATGGAAATTTGTTCA 1200
QY 1201 TAATATTAAAGAAAGAGCAGGGTTCAAGAGAGCCATGGGGCACTACGCTTTAGCTGAAC 1260
DB 1201 TAATATTAAAGAAAGAGCAGGGTTCAAGAGAGCCATGGGGCACTACGCTTTAGCTGAAC 1260

Qy	1261	GTGAGTGCACGCGAGCTCTGAGACTCACTGTGGGAGAGAGACAACCTAGAGACTCAAG	1320
Db	1261	GTGAGTGCACGCGAGCTCTGAGACTCACTGTGGGAGAGAGACAACCTAGAGACTCAAG	1320
Qy	1321	AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAT	1380
Db	1321	AGGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAT	1380
Qy	1381	TGCTGCACGAACCTCCTTGATTTTAGCCTTCTCTGTTCAATTCCTCAAAAAGATTCCCCA	1440
Db	1381	TGCTGCACGAACCTCCTTGATTTTAGCCTTCTCTGTTCAATTCCTCAAAAAGATTCCCCA	1440

RESULT 11

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```
QY 721 TGGAGTGGGAAGGCAACAAGATTTCGGCCAGGACAGCAACAAGGCGCTACCTGGAGAGGACT 780
Db 721 TGGAGTGGGAAGGCAACAAGATTTCGGCCAGGACAGCAACAAGGCGCTACCTGGAGAGGACT 780
QY 781 GCCTTCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
Db 781 GCCTTCACAGCTGCAGCAGTGTCTGGAGCTGGGAGAGGTGTTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGTTGAAGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCCGG 900
Db 841 CTCCTTTGTTGAAGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTCCGG 900
QY 901 CCTTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGATAAGACAGCAATGG 960
Db 901 CCTTGAACCTACTACCCCAAGACATCACCATTGAAGTGGCTGAAGATAAGACAGCAATGG 960
QY 961 ATGCCAAGAGTTGCGAACCTTAAGACGTATTCGCCAATGGGGATGGGACCTTACCAGGCT 1020
Db 961 ATGCCAAGAGTTGCGAACCTTAAGACGTATTCGCCAATGGGGATGGGACCTTACCAGGCT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGAGAGCAGAGATATACGTACCAGTGGAGACC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGAGAGCAGAGATATACGTACCAGTGGAGACC 1080
QY 1081 CAGGCTGGATCAGCCCTCATTTGATCTGGAGCCCTCAACCCTCTGGCACCTTAGTCA 1140
Db 1081 CAGGCTGGATCAGCCCTCATTTGATCTGGAGCCCTCAACCCTCTGGCACCTTAGTCA 1140
QY 1141 TTGGAGTCAATCAGTGAATGCTGTTTTTGTGCTCATCTTGTTCATTGGAATTTGTTCA 1200
Db 1141 TTGGAGTCAATCAGTGAATGCTGTTTTTGTGCTCATCTTGTTCATTGGAATTTGTTCA 1200
QY 1201 TAATATTAGGAAGAGGAGGCTTCAAGAGGAGCCATGGGCACTTACGCTTACGTGAAC 1260
Db 1201 TAATATTAGGAAGAGGAGGCTTCAAGAGGAGCCATGGGCACTTACGCTTACGTGAAC 1260
QY 1261 GTGAGTGACACGACCTGCAGACTACTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
Db 1261 GTGAGTGACACGACCTGCAGACTACTGTGGGAAGGAGACAAACTAGAGACTCAAAG 1320
QY 1321 AGGGAGTGCAATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTTAAACATAGAAAT 1380
Db 1321 AGGGAGTGCAATTATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACCTTAAACATAGAAAT 1380
QY 1381 TGCCTGACAGCACTCTTGATTTAGCTTCTCTGTTTCATTTTCCCTCAAAAAGATTTCCCCA 1440
Db 1381 TGCCTGACAGCACTCTTGATTTAGCTTCTCTGTTTCATTTTCCCTCAAAAAGATTTCCCCA 1440
```

```
RESULT 12
US-09-503-444A-9
; Sequence 9, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/652,265
; FILING DATE: 23-May-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,673
; FILING DATE: 16-Apr-1996
; APPLICATION NUMBER: 08/630,912
; FILING DATE: 04-Apr-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0088-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1440 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 222..1268
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(414, "a")
; OTHER INFORMATION: /phenotype= "normal or wild-type"
; OTHER INFORMATION: (unaffected)
; OTHER INFORMATION: /label= 24d7
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(1066, "g")
; OTHER INFORMATION: /phenotype= "normal or wild-type"
; OTHER INFORMATION: (unaffected)
; OTHER INFORMATION: /label= 24d1
; US-09-503-444A-9
```

```
Query Match 99.8%; Score 1436.8; DB 4; Length 1440;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGACACCTGGATCAGCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
Db 1 GGGACACCTGGATCAGCTAGTGTTCACAAGCAGGTACCTTCTGCTGTAGGAGAGAGA 60
QY 61 ACTAAAGTTCTGAAAGACCTGTGCTTTTCCACAGGAAGTTTACTGGGCATCTCTCTGAG 120
Db 61 ACTAAAGTTCTGAAAGACCTGTGCTTTTCCACAGGAAGTTTACTGGGCATCTCTCTGAG 120
QY 121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCAAAAG 180
Db 121 CCTAGGCAATAGCTGTAGGGTGACTTCTGGAGCCATCCCCGTTTCCCCGCCCCCAAAAG 180
QY 181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGCCCGCGAGCCAGGC 240
Db 181 AAGCGGAGATTTAACGGGGACGTGCGGCCAGAGCTGGGGAATGGGCCCGCGAGCCAGGC 240
```

QY 241 CGCGCTTCTCCTCGATGCTTTTGCAGACCGCGTCTCTCAGGGGCGCTTCTCGCGTT 300
Db 241 CGCGCTTCTCCTCGATGCTTTTGCAGACCGCGTCTCTCAGGGGCGCTTCTCGCGTT 300
QY 301 CACACTCTCTGCACACTACCTTCTCATGGGTGCCTCAGAGCAGACCTTGCTCTTCTTGT 360
Db 301 CACACTCTCTGCACACTACCTTCTCATGGGTGCCTCAGAGCAGACCTTGCTCTTCTTGT 360
QY 361 TTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCGTCTTATGATGAGATCGCC 420
Db 361 TTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTCGTCTTATGATGAGATCGCC 420
QY 421 GTCTGAGCCCCGAACCTCCATGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
Db 421 GTCTGAGCCCCGAACCTCCATGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAGC 480
QY 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTCTTGACTTCTGGACTATTATGG 540
Db 481 TGAGTCAGAGTCTGAAGGGTGGGATCACATGTTCACTCTTGACTTCTGGACTATTATGG 540
QY 541 AAAATCACAAACACAGAGGAGTCCACACCCCTGCAGGTCTATCCTGGGCTGTGAATGC 600
Db 541 AAAATCACAAACACAGAGGAGTCCACACCCCTGCAGGTCTATCCTGGGCTGTGAATGC 600
QY 601 AAGAAGCACACACTACCGAGGGCTACTGGAAGTACGGGTATGATGGGACGACCCTTG 660
Db 601 AAGAAGCACACACTACCGAGGGCTACTGGAAGTACGGGTATGATGGGACGACCCTTG 660
QY 661 AATTCTGCCCTGACACACTGGAATTTGAGAGCAGACGCCAGGGCTGGCCCCACCAAGC 720
Db 661 AATTCTGCCCTGACACACTGGAATTTGAGAGCAGACGCCAGGGCTGGCCCCACCAAGC 720
QY 721 TGGAGTGGGAAGGCCACAAGATTCGGGCCAGGCAGAACAGGGCCCTACCTGGAGGGACT 780
Db 721 TGGAGTGGGAAGGCCACAAGATTCGGGCCAGGCAGAACAGGGCCCTACCTGGAGGGACT 780
QY 781 GCCCTGCACAGCTGCAGAGTGTCTGGAGCTGGGAGAGGTGTTTTGGACCAACAAGTGC 840
Db 781 GCCCTGCACAGCTGCAGAGTGTCTGGAGCTGGGAGAGGTGTTTTGGACCAACAAGTGC 840
QY 841 CTCCTTTGTTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTCTACGGTGTGCGG 900
Db 841 CTCCTTTGTTGAAGGTGACACATCATGTGACCTCTTCACTGACCACTCTACGGTGTGCGG 900
QY 901 CCTTGAACACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGATGAAGCAACCAATGG 960
Db 901 CCTTGAACACTACTACCCCCAGAACATCACCATGAAGTGGCTGAAGATGAAGCAACCAATGG 960
QY 961 ATCCCAAGGAGTTCGAACCTAAAGAGGTATTGCCCAATGGGATGGGACCTACCGGGT 1020
Db 961 ATCCCAAGGAGTTCGAACCTAAAGAGGTATTGCCCAATGGGATGGGACCTACCGGGT 1020
QY 1021 GGATAACCTTGGCTGTACCCCTGGGAGAGACAGATATACGTACAGGTGGAGCACC 1080
Db 1021 GGATAACCTTGGCTGTACCCCTGGGAGAGACAGATATACGTGCGCAGGTGGAGCACC 1080
QY 1081 CAGGCTGGATCAGCCCTCATTGTGATCTGGAGCCCTCACCGCTGTGGCACCTAGTCA 1140
Db 1081 CAGGCTGGATCAGCCCTCATTGTGATCTGGAGCCCTCACCGCTGTGGCACCTAGTCA 1140
QY 1141 TTGAGTCTACAGTGAATTGTGTTTTTGTCTCATCTTGTTCATTTGGAATTTTGTTC 1200
Db 1141 TTGAGTCTACAGTGAATTGTGTTTTTGTCTCATCTTGTTCATTTGGAATTTTGTTC 1200
QY 1201 TAATATTAGGAGAGGCGAGGTTCAAGAGAGCCATGGGACCTACGCTTTAGCTGAAC 1260
Db 1201 TAATATTAGGAGAGGCGAGGTTCAAGAGAGCCATGGGACCTACGCTTTAGCTGAAC 1260
QY 1261 GTGAGTGACACGCGCTGCAGACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320
Db 1261 GTGAGTGACACGCGCTGCAGACTCAGTGTGGGAAGGAGACAAACTAGAGACTCAAG 1320

QY 1321 AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
Db 1321 AGGAGTGCAATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTAAACATAGAAAT 1380
QY 1381 TGCCTGAGAACTCCTTGATTTAGCCTTCTCTCATTTCCCAAAAAGATTTCGCCA 1440
Db 1381 TGCCTGAGAACTCCTTGATTTAGCCTTCTCTCATTTCCCAAAAAGATTTCGCCA 1440
RESULT 13
US-09-277-457-1
; Sequence 1, Application US/09277457
; Patent No. 635425
; GENERAL INFORMATION:
; APPLICANT: Rothenberg, Barry E.
; APPLICANT: Sawada-Hirai, Ritsuko
; APPLICANT: Barton, James C.
; TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
; FILE REFERENCE: 10653/002001
; CURRENT APPLICATION NUMBER: US/09/277,457
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2506
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: Missense mutation at nucleotide 314
US-09-277-457-1
Query Match 84.4%; Score 1215,8; DB 4; Length 2506;
Best Local Similarity 99,8%; Pred. No. 0;
Matches 1217; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 222 ATGGGCCCGCAGCGCGCTTCTCCTCGATGCTTTTGCAGACCGCGTCTG 281
Db 1 atgggcccgcgagccaggccgcttctctctctctctctctctctctctctctctctg 60
QY 282 CAGGGGCGCTTGTGCGTTTCACACTCTCTGCTGCTACCTTCTCATGGTGGCTCAGACG 341
Db 61 cagggcgctgtgcgttcacactctctgcactacctcttcctcatgggtgcctcagacg 120
QY 342 GACCTTGTCTTCTCCTTGTAAAGCTTTGGGCTTACGTGGATGACCACTGTTCTGCTGTC 401
Db 121 gaccttggctcttctctgtttgaagctttgggtacgtggatgaccagctgttcgtgttc 180
QY 402 TATGATGATGAGCTCGCGCTGTGGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCA 461
Db 181 tatgatcatgagagtcgctgtgtgagcccgaaactccatgggtttccagtagaatcca 240
QY 462 AGCCAGATGTGGCTGACGCTCAGAGTCTGAAAGGTGGGATCACATGTTTCACTGTT 521
Db 241 agccagatgtggctgacgctcagagtcagagtcgaaaggtgggatacacaigtacactgtt 300
QY 522 GACTTCTGGACTATTATGAAATCACAACACAGAGGAGTCCACACCCCTCGAGGTC 581
Db 301 gactctggactattatgaaatacacaacacagagagtgagtgccacacccctcgaggtc 360
QY 582 ATCTTGGGCTGTAAATGCAAGAAACAACTACCGAGGGCTACTGGAAGTACGGGTAT 641
Db 361 atcttgggctgtgaaatgcaagaagacaacagtagccagggctactggaagtacgggtat 420
QY 642 GATGGCAGGACCACTTGAATTCGCCCTTGACACACTGGATTTGGAGCAGCAGAACCC 701
Db 421 gatggcagggaccaccttgaattctgccttgacacactggattgagagcagcagaacc 480
QY 702 AGGCGCTGGCCCCACCAAGCTGGAGTGGAAAGGCACAAAGATTTCGGGCCAGGCACAGG 761
Db 481 agggcctggcccccaagctggagtgggaaaggcacaagattcggggccagcagaacagg 540

Db 3942 AGCTGAGTCAGAGTCGGAAGGTTGGGATCACATGTTCACTGTTGACTTCTTGGACTATTA 4001
QY 538 TGGAAATACACACACACAG- 560
Db 4002 TGGAAATACACACACACAGAGGTTATGTGGAGAGGGCCCTCACCTTCTGAGGTTGT 4061
QY 561 ----- 560
Db 4062 CAGAGCTTTTTCATCTTTTCATGTCATCTTGAAGGAACAGCTGGAAGTCTGAGGCTTTGTG 4121
QY 561 ----- 560
Db 4122 GGAGCAGGGAAGAGGAATTTGCTTCTCCTGAGATCATTTGGTCTTGGGATGGTG 4181
QY 561 -----GAGTCCCA 568
Db 4182 AATAGGAGCTTATCTTCTTGGTTGCAGTTAACAAGCTGGGATTTTCCACAGTCCCA 4241
QY 569 CACCTGTCAGGTCATCTGCGCTGTGAATGCAAGAGACACAGTACCGAGGCTACTG 628
Db 4242 CACCTGTCAGGTCATCTGCGCTGTGAATGCAAGAGACACAGTACCGAGGCTACTG 4301
QY 629 GAAGTACGGTATGATGGCAGACACACCTTGAATCTGCCCTGACACACTGGATTGGAG 688
Db 4302 GAAGTACGGTATGATGGCAGACACCTTGAATCTGCCCTGACACACTGGATTGGAG 4361
QY 689 AGCAGCAGAACCCAGGCGCTGGCCCAACCAAGCTGGAGTGGGAAGGACACAAAGATTGGGC 748
Db 4362 AGCAGCAGAACCCAGGCGCTGGCCCAACCAAGCTGGAGTGGGAAGGACACAAAGATTGGGC 4421
QY 749 CAGCAGACAGGCGCTACTGTCAGAGGAGTGCCTCTGACACACTGACAGAGTTGCTGGA 808
Db 4422 CAGCAGACAGGCGCTACTGTCAGAGGAGTGCCTCTGACACACTGACAGAGTTGCTGGA 4481
QY 809 GCTGGGAGAGGTTTGGACCAACAG 837
Db 4482 GCTGGGAGAGGTTTGGACCAACAG 4510

RESULT 15
US-08-652-265-7
Sequence 7, Application US/08652265
Patent No. 6025130
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,265
FILING DATE: 23-MAY-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223

REFERENCE/DOCKET NUMBER: 17957-000500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: and 24d2 mutations"
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)"
OTHER INFORMATION: gene containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles"
FEATURE:
NAME/KEY:
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: cDNA containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles"
OTHER INFORMATION:
FEATURE:
NAME/KEY:
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
FEATURE:
NAME/KEY:
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d1
US-08-652-265-7
Query Match 22.3%; Score 321; DB 3; Length 10825;
Best Local Similarity 72.1%; Pred. No. 4.8e-83;
Matches 540; Conservative 0; Mismatches 0; Indels 209; Gaps 1;
QY 298 GTTCACACTCTCTGCACACTACCTCTTCATGGGTGCCTCAGACGAGACCTTGGTCTTTCCCT 357
Db 3762 GTTCACACTCTCTGCACACTACCTCTTCATGGGTGCCTCAGACGAGACCTTGGTCTTTCCCT 3821
QY 358 TGTTTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTGCTGTTCTATGATGATGAGATC 417
Db 3822 TGTTTGAAGCTTTGGGCTACGTGGATGACACAGCTGTTGCTGTTCTATGATGATGAGATC 3881
QY 418 GCCGTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 477
Db 3882 GCCGTGTGGAGCCCGCAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGC 3941
QY 478 AGCTGAGTCAGAGTCTGAAAGGGTGGGATCATGTTCACTGTTGACTTCTGGACTATTA 537

Db 3942 AGCTGAGTCAGAGTCTGAAGGGTGGGATCACAATGTTCACTGTGACTTCTGGACTATTA 4001
QY 538 TGGAAATCACAAACACAGCAAG----- 560
|||||
Db 4002 TGGAAATCACAAACACAGAGGTATGTGGAGAGGGGCTCACCTTCTGAGTTGT 4061
QY 561 ----- 560
Db 4062 CAGAGCTTTTCATCTTTTCATGATCTTGAAGGAACAGCTGGAAGTCTGAGTCTTTGG 4121
QY 561 ----- 560
Db 4122 GGAGCAGGAAGAGGAAGGAATTTGCTTCTGAGATCATTTGGTCTTGGGATGGTGG 4181
QY 561 -----GAGTCCCA 568
|||||
Db 4182 AAATAGGACCTATTCTCTTTGTTGAGTTAACAAGGCTGGGATTTTCCAGAGTCCCA 4241
QY 569 CACCTTGCAGGTATCTCTGGGCTGTGAATGCAAGAAGACACAGTACCGAGGGCTACTG 628
|||||
Db 4242 CACCTTGCAGGTATCTCTGGGCTGTGAATGCAAGAAGACACAGTACCGAGGGCTACTG 4301
QY 629 GAAGTACGGGTATGATGGGCAGGACACCTTGAATTTGCCCTGACACACTGGATTGGAG 688
|||||
Db 4302 GAAGTACGGGTATGATGGGCAGGACACCTTGAATTTGCCCTGACACACTGGATTGGAG 4361
QY 689 AGCAGCAGAACCCAGGGCTTGCCCAACCAAGCTGGAGTGGGAAAGGCACAAAGATTCGGGC 748
|||||
Db 4362 AGCAGCAGAACCCAGGGCTTGCCCAACCAAGCTGGAGTGGGAAAGGCACAAAGATTCGGGC 4421
QY 749 CAGGCAGAACAGGGCTTACCTGGAGAGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 808
|||||
Db 4422 CAGGCAGAACAGGGCTTACCTGGAGAGGACTGCCCTGCACAGCTGCAGCAGTTGCTGGA 4481
QY 809 GCTGGGGAGAGGTGTTTGGACCAACAAG 837
|||||
Db 4482 GCTGGGGAGAGGTGTTTGGACCAACAAG 4510

Search completed: June 19, 2002, 08:40:57
Job time: 7342 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:40:02 ; Search time 1828.83 Seconds
(without alignments)
3444.217 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870
Perfect score: 301
Sequence: 1 GGCACGATCCCTGTTGG.....GCACTTGGAGCGCGAGGC 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

ALIGNMENTS

RESULT 1
AR117789
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
BASE COUNT
ORIGIN

AR117789
Sequence
AR117789
AR117789.1
Unknown.
Unclassified.
Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
Hereditary hemochromatosis gene products
Patent: US 6140305-A 1 31-OCT-2000;
Location/Qualifiers
1..10825
/organism="unknown"
2998 a 2253 c 2648 g 2926 t

10825 bp
1 from patent US 6140305.
DNA
linear
PAT 16-MAY-2001

AR117789 Sequence
AR117790 Sequence
AR117791 Sequence
AR117792 Sequence
AR149459 Sequence
AR149460 Sequence
AR149461 Sequence
AR149462 Sequence
292910 Homo sapien
AR036572 Sequence
AR036573 Sequence
AR036574 Sequence
U91328 Human hered
AL359892 Homo sapi
AF301592 Dicerob b
AC055888 Homo sapi
AC073212 Homo sapi
AC093333 Homo sapi
AC000385 Homo sapi
AL360004 Human DNA
AC060833 Homo sapi
AC015906 Homo sapi
AC027062 Homo sapi
AC018362 Homo sapi
AC016605 Homo sapi
AF111168 Homo sapi
AJ009615 Homo sapi
AC069317 Homo sapi
AC018464 Homo sapi
AC015922 Homo sapi
AL354046 Homo sapi
AC013617 Homo sapi
AC087315 Homo sapi
AC016555 Homo sapi
AC084116 Homo sapi
AC010897 Homo sapi
AL353729 Human DNA
AC021396 Homo sapi
AL591222 Homo sapi
AC013475 Homo sapi
AL157373 Human DNA
AC008917 Homo sapi
AP003496 Homo sapi
AC023590 Homo sapi
AL137073 Human DNA

Query Match 100.0%; Score 301; DB 6; Length 10825;

```
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

Qy 61 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 4689

Qy 121 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749

Qy 181 GCCTTGCTTTTATTTACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCTTGCTTTTATTTACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 4809

Qy 241 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 300
Db 4810 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 4869

Qy 301 C 301
Db 4870 C 4870

RESULT 2
AR117790 AR117790 10825 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 3 from patent US 6140305.
DEFINITION AR117790
ACCESSION AR117790
VERSION AR117790.1 GI:14098696
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuhiihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 3 31-OCT-2000;
FEATURES
source
1..10825
/organism="unknown"
BASE COUNT 2999 a 2253 c 2647 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

Qy 61 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 4689

Qy 121 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749

Qy 181 GCCTTGCTTTTATTTACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCTTGCTTTTATTTACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 4809

Qy 241 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 300
Db 4810 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 4869

Qy 301 C 301
Db 4870 C 4870

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

Qy 61 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 4689

Qy 121 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749

Qy 181 GCCTTGCTTTTATTTACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCTTGCTTTTATTTACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 4809

Qy 241 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 300
Db 4810 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 4869

Qy 301 C 301
Db 4870 C 4870
```

```
Qy 301 C 301
Db 4870 C 4870

RESULT 3
AR117791 AR117791 10825 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 5 from patent US 6140305.
DEFINITION AR117791
ACCESSION AR117791
VERSION AR117791.1 GI:14098697
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuhiihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 5 31-OCT-2000;
FEATURES
source
1..10825
/organism="unknown"
BASE COUNT 2998 a 2252 c 2649 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTGGAGCTTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

Qy 61 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGACATCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 4689

Qy 121 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTAAATTTCTTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749

Qy 181 GCCTTGCTTTTATTTAAACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCTTGCTTTTATTTAAACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGA 4809

Qy 241 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 300
Db 4810 AATGTCAGGCGGGGACGGTGGCTGCCTCAACCCCTGTAATCCAGCACTTTGGAGGCCGAGG 4869

Qy 301 C 301
Db 4870 C 4870

RESULT 4
AR117792 AR117792 10825 bp DNA linear PAT 16-MAY-2001
LOCUS Sequence 7 from patent US 6140305.
DEFINITION AR117792
ACCESSION AR117792
VERSION AR117792.1 GI:14098698
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuhiihashi,Z. and Wolff,R.K.
TITLE Hereditary hemochromatosis gene products
JOURNAL Patent: US 6140305-A 7 31-OCT-2000;
FEATURES
source
1..10825
/organism="unknown"
BASE COUNT 2998 a 2252 c 2649 g 2926 t
ORIGIN
```

```
source 1. 10825
/organism="unknown"
BASE COUNT 2999 a 2252 c 2648 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

QY 61 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTAATAATTGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTAATAATTGAGATGATGAGACAGC 4689

QY 121 CACAAGTCATGGGTTTAATTTCTTTCCTCATGATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTTAATTTCTTTCCTCATGATGGCTCAAGGGAAGTGTCTATG 4749

QY 181 GCCCTTGCCTTTTATTTAACCAATAAATCTTTGTATATTTATACCTGTATAAAATTCAGA 240
Db 4750 GCCCTTGCCTTTTATTTAACCAATAAATCTTTGTATATTTATACCTGTATAAAATTCAGA 4809

QY 241 ATGTGCAAGGCGGCGACGGTGGCTCACCCTGTATATCCAGCAGCTTTGGGAGGCGGAGG 300
Db 4810 ATGTGCAAGGCGGCGACGGTGGCTCACCCTGTATATCCAGCAGCTTTGGGAGGCGGAGG 4869

QY 301 C 301
Db 4870 C 4870

RESULT 5
LOCUS ARI49459 10825 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6228594.
ACCESSION ARI49459
VERSION ARI49459.1 GI:15114050
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuchihashi,Z. and Wolff,R.K.
TITLE Method for determining the presence or absence of a hereditary
hemochromatosis gene mutation
JOURNAL Patent: US 6228594-A 1 08-MAY-2001;
FEATURES
source
location/Qualifiers
1. 10825
/organism="unknown"
BASE COUNT 2998 a 2253 c 2648 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

QY 61 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTAATAATTGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTAATAATTGAGATGATGAGACAGC 4689

QY 121 CACAAGTCATGGGTTTAATTTCTTTCCTCATGATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTTAATTTCTTTCCTCATGATGGCTCAAGGGAAGTGTCTATG 4749
```

```
QY 181 GCCCTTGCCTTTTATTTAACCAATAAATCTTTGTATATTTATACCTGTATAAAATTCAGA 240
Db 4750 GCCCTTGCCTTTTATTTAACCAATAAATCTTTGTATATTTATACCTGTATAAAATTCAGA 4809

QY 241 ATGTGCAAGGCGGCGACGGTGGCTCACCCTGTATATCCAGCAGCTTTGGGAGGCGGAGG 300
Db 4810 ATGTGCAAGGCGGCGACGGTGGCTCACCCTGTATATCCAGCAGCTTTGGGAGGCGGAGG 4869

QY 301 C 301
Db 4870 C 4870

RESULT 6
LOCUS ARI49460 10825 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 3 from patent US 6228594.
ACCESSION ARI49460
VERSION ARI49460.1 GI:15114051
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D.,
Tsuchihashi,Z. and Wolff,R.K.
TITLE Method for determining the presence or absence of a hereditary
hemochromatosis gene mutation
JOURNAL Patent: US 6228594-A 3 08-MAY-2001;
FEATURES
source
location/Qualifiers
1. 10825
/organism="unknown"
BASE COUNT 2999 a 2253 c 2647 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

QY 61 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTAATAATTGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTAATAATTGAGATGATGAGACAGC 4689

QY 121 CACAAGTCATGGGTTTAATTTCTTTCCTCATGATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTTAATTTCTTTCCTCATGATGGCTCAAGGGAAGTGTCTATG 4749

QY 181 GCCCTTGCCTTTTATTTAACCAATAAATCTTTGTATATTTATACCTGTATAAAATTCAGA 240
Db 4750 GCCCTTGCCTTTTATTTAACCAATAAATCTTTGTATATTTATACCTGTATAAAATTCAGA 4809

QY 241 ATGTGCAAGGCGGCGACGGTGGCTCACCCTGTATATCCAGCAGCTTTGGGAGGCGGAGG 300
Db 4810 ATGTGCAAGGCGGCGACGGTGGCTCACCCTGTATATCCAGCAGCTTTGGGAGGCGGAGG 4869

QY 301 C 301
Db 4870 C 4870

RESULT 7
LOCUS ARI49461 10825 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 5 from patent US 6228594.
ACCESSION ARI49461
VERSION ARI49461.1 GI:15114052
KEYWORDS
```

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Method for determining the presence or absence of a hereditary hemochromatosis gene mutation
JOURNAL Patent: US 6228594-A 5 08-MAY-2001;
FEATURES Location/Qualifiers
 1..10825
 /organism="unknown"
BASE COUNT 2998 a 2252 c 2649 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGCTCTCCAAATTC 60
DB 4570 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGCTCTCCAAATTC 4629

QY 61 TGGGAAGGACCTTCATCCTAGAGTCTACCTTATAATAGATGTATGAGACAGC 120
DB 4630 TGGGAAGGACCTTCATCCTAGAGTCTACCTTATAATAGATGTATGAGACAGC 4689

QY 121 CACAAGTCATGGTTTAAATTTCTTCCATGCATATGGCTCAAAGGAAGTGTCTATG 180
DB 4690 CACAAGTCATGGTTTAAATTTCTTCCATGCATATGGCTCAAAGGAAGTGTCTATG 4749

QY 181 GCCCTTGCTTTTATTAAACCAATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 240
DB 4750 GCCCTTGCTTTTATTAAACCAATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

QY 241 AATGTCAGGCGGCGGACGGTGGCTCACCCCTGTAATCCAGCACCTTTGGGAGCGCGAGG 300
DB 4810 AATGTCAGGCGGCGGACGGTGGCTCACCCCTGTAATCCAGCACCTTTGGGAGCGCGAGG 4869

QY 301 C 301
DB 4870 C 4870

RESULT 8
LOCUS ARI49462
DEFINITION Sequence 7 from patent US 6228594.
ACCESSION ARI49462
VERSION ARI49462.1 GI:15114053
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 10825)
AUTHORS Thomas,W.J., Drayna,D.T., Feder,J.N., Gnirke,A., Ruddy,D., Tsuchihashi,Z. and Wolff,R.K.
TITLE Method for determining the presence or absence of a hereditary hemochromatosis gene mutation
JOURNAL Patent: US 6228594-A 7 08-MAY-2001;
FEATURES Location/Qualifiers
 1..10825
 /organism="unknown"
BASE COUNT 2999 a 2252 c 2648 g 2926 t
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGCTCTCCAAATTC 60

DB 4570 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGCTCTCCAAATTC 4629

QY 61 TGGGAAGGACCTTCATCCTAGAGTCTACCTTATAATAGATGTATGAGACAGC 120

DB 4630 TGGGAAGGACCTTCATCCTAGAGTCTACCTTATAATAGATGTATGAGACAGC 4689

QY 121 CACAAGTCATGGTTTAAATTTCTTCCATGCATATGGCTCAAAGGAAGTGTCTATG 180

DB 4690 CACAAGTCATGGTTTAAATTTCTTCCATGCATATGGCTCAAAGGAAGTGTCTATG 4749

QY 181 GCCCTTGCTTTTATTAAACCAATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 240

DB 4750 GCCCTTGCTTTTATTAAACCAATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

QY 241 AATGTCAGGCGGCGGACGGTGGCTCACCCCTGTAATCCAGCACCTTTGGGAGCGCGAGG 300

DB 4810 AATGTCAGGCGGCGGACGGTGGCTCACCCCTGTAATCCAGCACCTTTGGGAGCGCGAGG 4869

QY 301 C 301

DB 4870 C 4870

RESULT 9
LOCUS HSHFE
DEFINITION Homo sapiens HFE gene.
ACCESSION 292910
VERSION 292910.1 GI:1890179
KEYWORDS haemochromatosis; HFE gene.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 858)
AUTHORS Albig,W., Drabent,B., Burmester,N., Bode,C. and Doenecke,D.
TITLE The haemochromatosis candidate gene HFE (HLA-H) of man and mouse is located in syntenic regions within the histone gene cluster
JOURNAL J. Cell. Biochem. 69 (2), 117-126 (1998)
MEDLINE 98208340
REFERENCE 2 (bases 1 to 12146)
AUTHORS Albig,W.
TITLE Direct Submission
JOURNAL Submitted (14-MAR-1997) Albig W., Georg-August-Universitaet Goettingen, Biochemie und Molekulare Zellbiologie, Humboldtallee 23, Goettingen, FRG, 37073
FEATURES Location/Qualifiers
 1..12146
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="6"
 /map="6p"
 /clone_lib="ICRF YAC-library"
 /clone="ICRFy901D1223"
 1028..1324
 /genes="HFE"
 /number=1
 1028..10637
 /genes="HFE"
 join(1249..1324,4652..4915,5125..5400,6494..6769,6928..7041,7995..8035)
 /genes="HFE"
 /function="iron metabolism"
 /notes="haemochromatosis candidate gene"
 /codon_start=1
 /protein_id="CAB07442.1"
 /db_xref="GI:1890180"
 /translation="MGPRAPALLLLMLLQTVLOGRLLRSLHSLHYLFMGASEQDLGLSFEALGYVDDQLFVFYDHESRRVETPWSSRISOMQLQSLKSGMDHMTVDFSLTTMENHNSKESHTLQVILGCEMDQEDNSTEGYKYGQDHLFCFPTLDWRAAEPRAWPTKLEWHRHKIRARQNRAYLERDCPAQLQQLLELGRGLDQQVPPVLKVTHTVTS

SVTLRCRALNYQPNTMKWLKQPMDAKEFEKPKDVLPGNGDTYQGWITLAVPPGE
EORTYCGVHPGLDOPDIVIWEPSPTLVIGVISIAVFWILFTIGILFIILKRQGG
SRGAMGHVLAERE"

intron 1325..4651
/gene="HFE"
/number=1
repeat_unit 3494..3735
/gene="HFE"
/rpt_family="Alu"
repeat_unit 3973..4283
/gene="HFE"
/rpt_family="Alu"
exon 4652..4915
/gene="HFE"

intron 4916..5124
/gene="HFE"
exon 5125..5400
/gene="HFE"
intron 5401..6493
/gene="HFE"

repeat_unit 5707..6005
/gene="HFE"
/rpt_family="Alu"
exon 6494..6769
/gene="HFE"
intron 6770..6927
/gene="HFE"
exon 6928..7041
/gene="HFE"
intron 7042..7794
/gene="HFE"
exon 7995..9050
/gene="HFE"

repeat_unit 9017..9340
/gene="HFE"
intron 9051..10205
/gene="HFE"
repeat_unit 9957..10239
/gene="HFE"
exon 10206..10637
/gene="HFE"
polya_signal 10617..10622
/gene="HFE"

BASE COUNT 3383 a 2474 c 2911 g 3378 t
ORIGIN

Query Match 100.0%; Score 301; DB 9; Length 12146;
Best Local Similarity 100.0%; Pred. No. 1.2e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCCTCCAAATTC 60
DB 5460 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCCTCCAAATTC 5519
QY 61 TGGGAAGGGACTTTCATCTAGAGTCTCTACCTTATAATAGATGATGAGACAGC 120
DB 5520 TGGGAAGGGACTTTCATCTAGAGTCTCTACCTTATAATAGATGATGAGACAGC 5579
QY 121 CACAAGTCATGGGTTTAATTTCTTTCTCCATCATATGGCTCAAGGGAAGTGCTATG 180

Db 5580 CACAAGTCATGGGTTTAATTTCTTTCTCCATCATATGGCTCAAGGGAAGTGCTATG 5639
QY 181 GCCCTGGCTTTTATTAACCAATAATCTTTGTATATTTATACCTGTAAATAATTCAGA 240
DB 5640 GCCCTGGCTTTTATTAACCAATAATCTTTGTATATTTATACCTGTAAATAATTCAGA 5699
QY 241 AATGTCAAGCGCGGCAGCGTGGCTCACCCCTGTAATCCAGCACATTTGGGAGCCGAGG 300
DB 5700 AATGTCAAGCGCGGCAGCGTGGCTCACCCCTGTAATCCAGCACATTTGGGAGCCGAGG 5759
QY 301 C 301
DB 5760 C 5760

RESULT 10
AR036572 AR036572 246240 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 20 from patent US 5872237.
DEFINITION AR036572
ACCESSION AR036572
VERSION AR036572.1 GI:5953240
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 246240)
AUTHORS Feder, J. Nathan., Kronmal, G. Scott., Lauer, P. M., Ruddy, D. A.,
Thomas, W., Tsuchibashi, Z. and Wolff, R. K.
TITLE Megabase transcript map: novel sequences and antibodies thereto
JOURNAL Patent: US 5872237-A 20 16-FEB-1999;
FEATURES Location/Qualifiers
source 1..246240
BASE COUNT 73211 a 50177 c 50599 g 72252 t 1 others
ORIGIN

Query Match 100.0%; Score 301; DB 6; Length 246240;
Best Local Similarity 100.0%; Pred. No. 1.4e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCCTCCAAATTC 60
DB 196873 GGCACGGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCCTCCAAATTC 196932
QY 61 TGGGAAGGGACTTTCATCTAGAGTCTCTACCTTATAATGAGATGATGAGACAGC 120
DB 196933 TGGGAAGGGACTTTCATCTAGAGTCTCTACCTTATAATGAGATGATGAGACAGC 196992
QY 121 CACAAGTCATGGGTTTAATTTCTTTCTCCATCATATGGCTCAAGGGAAGTGCTATG 180
DB 196993 CACAAGTCATGGGTTTAATTTCTTTCTCCATCATATGGCTCAAGGGAAGTGCTATG 197052
QY 181 GCCCTGGCTTTTATTAACCAATAATCTTTGTATATTTATACCTGTAAATAATTCAGA 240
DB 197053 GCCCTGGCTTTTATTAACCAATAATCTTTGTATATTTATACCTGTAAATAATTCAGA 197112
QY 241 AATGTCAAGCGCGGCAGCGTGGCTCACCCCTGTAATCCAGCACATTTGGGAGCCGAGG 300
DB 197113 AATGTCAAGCGCGGCAGCGTGGCTCACCCCTGTAATCCAGCACATTTGGGAGCCGAGG 197172
QY 301 C 301
DB 197173 C 197173

RESULT 11
AR036573 AR036573 246240 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 21 from patent US 5872237.
DEFINITION AR036573
ACCESSION AR036573
VERSION AR036573.1 GI:5953241

repeat_region	8203. .8463 /rpt_family="Alu"	repeat_region	/rpt_type=dispersed 39132. .39252
repeat_region	8507. .8616 /rpt_family="MSTAR"	repeat_region	/rpt_type=dispersed complement(40045. .40645)
repeat_region	complement(8888. .9194) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed 40960. .41246
repeat_region	complement(9523. .9792) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed 41581. .42174
repeat_region	10109. .10375 /rpt_family="MSTAR"	repeat_region	/rpt_type=dispersed 42432. .42541
repeat_region	complement(10384. .10680) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed 44636. .44911
repeat_region	11482. .11742 /rpt_type=dispersed	repeat_region	/rpt_type=dispersed complement(45532. .45851)
repeat_region	complement(12899. .13326) /rpt_family="Alu"	gene	/rpt_type=dispersed complement(46833. .53618)
repeat_region	15021. .15309 /rpt_family="Alu"	CDS	/note="HFE" complement(join(46833. .46873,47827. .47940,48099. .48374, 49470. .49745,49955. .50218,53543. .53618))
repeat_region	complement(15450. .15714) /rpt_family="Alu"		/gene="HLA-H"
CDS	/rpt_type=dispersed complement(16434. .16826) /codon_start=1		/product="hereditary hemochromatosis"
	/product="histone 2A-like protein"		/protein_id="AAB82083.1"
	/db_xref="GI:2088554"		/db_xref="GI:2088551"
	GAGAPVLAAYLEVTAEILELAGNAARDNKTRIIPRHQLAIRNDEELKLLGRVT IAQGGVLPNTQAVLLPKTESHKAKGK"		/translation="MGPRARPALLLLMLLIQTAVLOGRLLSHLSHLVFMGASEQDLGL SLFEALGYVDDQLFYFHESRRVEPRTPWSSRISSQWNLQLSLKGWDMFTVDF WTIMNNHSHKESHTLVILGCEMQEDNSTEGYKVIYDQGHLEFCEDTLDNRAEP RAWPTKLEWERHKIRARQNAYLERDCPAQLQOLLELGRVLDQVPPPLVKTWHVTS SVTTLRCRALNTYPQNTMKWLKQPKMDAKEPEKDLVNGDGTQGMVTLVAPPG EORYTCQVHPGLDQPLIVIWEPSPGTLVIGVISGIAVVFVILFIIILRRQGG SRGAMGHVYLAERE"
repeat_region	complement(18988. .19282) /rpt_family="Alu"	repeat_region	complement(48868. .49182)
repeat_region	complement(20597. .20869) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed complement(50599. .50888)
repeat_region	complement(24343. .24624) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed complement(51145. .51372)
repeat_region	25654. .25805 /rpt_family="MIR"	repeat_region	/rpt_type=dispersed 51820. .51934
repeat_region	29842. .29938 /rpt_type=dispersed	repeat_region	/rpt_type=dispersed 55276. .55719
repeat_region	complement(30592. .30869) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed complement(56085. .56461)
repeat_region	31223. .31507 /rpt_type=dispersed	repeat_region	/rpt_family="L1" /rpt_type=dispersed complement(56521. .56801)
repeat_region	33810. .34150 /rpt_family="MER1"	repeat_region	/rpt_type=dispersed 57660. .57815
repeat_region	34235. .34314 /rpt_family="Alu"	repeat_region	/rpt_type=dispersed 58118. .58483
repeat_region	complement(35768. .36062) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed 58672. .58900
repeat_region	complement(36267. .36548) /rpt_family="Alu"	repeat_region	/rpt_type=dispersed complement(59299. .59737)
		repeat_region	/rpt_family="L1R2" complement(65665. .66013)


```

/rpt_family="LTR2"
/rpt_type=dispersed
66764. .67053
/rpt_family="Alu"
/rpt_type=dispersed
67743. .68016

Query Match      100.0%; Score 301; DB 9; Length 246282;
Best Local Similarity 100.0%; Pred. No. 1.4e-76;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGGATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 49410 GGCACGGATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 49351

QY 61 TGGGAAGGACATTCCTCAATCTAGAGTCTCTACCTTATAATGAGATGATGACAGC 120
Db 49350 TGGGAAGGACATTCCTCAATCTAGAGTCTCTACCTTATAATGAGATGATGACAGC 49291

QY 121 CACAAGTCATGGTTTAAATTTCTTTTCCATGTCATATGGCTCAAAGGAAAGTGTCTATG 180
Db 49290 CACAAGTCATGGTTTAAATTTCTTTTCCATGTCATATGGCTCAAAGGAAAGTGTCTATG 49231

QY 181 GCCCTTGCTTTTATTTAAACCAATAAATCTTTTGTATATTTATACCTGTTAAATTCAGA 240
Db 49230 GCCCTTGCTTTTATTTAAACCAATAAATCTTTTGTATATTTATACCTGTTAAATTCAGA 49171

QY 241 AATGTCAGGCGGCGAGGTGGCTGACCCCTGTAATCCAGCACACTTTGGAGGCCGAGG 300
Db 49170 AATGTCAGGCGGCGAGGTGGCTGACCCCTGTAATCCAGCACACTTTGGAGGCCGAGG 49111

QY 301 C 301
Db 49110 C 49110

RESULT 14
LOCUS AL359892/c
DEFINITION Homo sapiens chromosome 6 clone RP11-557F22, *** SEQUENCING IN
PROGRESS ***, 18 unordered pieces.
ACCESSION AL359892
VERSION AL359892.5 GI:9930971
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Sims.S.
Direct Submission
Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: Clonerequest@sanger.ac.uk
On Aug 27, 2000 this sequence version replaced gi:9864230.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA557F22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 183925 bases at least Q40
Consensus quality: 187703 bases at least Q30
Consensus quality: 189658 bases at least Q20
Insert size: 192052; sum-of-contigs
Insert size: 198247; agarose-fp
Quality coverage: 3.68x in Q20 bases; sum-of-contigs Quality
coverage: 3.70x in Q20 bases; agarose-fp

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3250: contig of 3250 bp in length
* 3251 3350: gap of 100 bp
* 3351 14600: contig of 11250 bp in length
* 14601 14700: gap of 100 bp
* 14701 32357: contig of 17657 bp in length
* 32358 32457: gap of 100 bp
* 32458 34886: contig of 2429 bp in length
* 34887 34986: gap of 100 bp
* 34987 43490: contig of 8504 bp in length
* 43491 43590: gap of 100 bp
* 43591 47437: contig of 3847 bp in length
* 47438 47537: gap of 100 bp
* 47538 57356: contig of 9819 bp in length
* 57357 57456: gap of 100 bp
* 57457 59845: contig of 2389 bp in length
* 59846 63972: contig of 4027 bp in length
* 63973 64072: gap of 100 bp
* 64073 82711: contig of 18639 bp in length
* 82712 82811: gap of 100 bp
* 82812 111814: contig of 29003 bp in length
* 111815 111914: gap of 100 bp
* 111915 120276: contig of 8362 bp in length
* 120277 120376: gap of 100 bp
* 120377 136660: contig of 16284 bp in length
* 136661 136760: gap of 100 bp
* 136761 153913: contig of 17153 bp in length
* 153914 154013: gap of 100 bp
* 154014 158659: contig of 4646 bp in length
* 158660 158759: gap of 100 bp
* 158760 164235: contig of 5476 bp in length
* 164236 164335: gap of 100 bp
* 164336 184996: contig of 20661 bp in length
* 184997 185096: gap of 100 bp
* 185097 193752: contig of 8656 bp in length.
*
* Location/Qualifiers
* 1..193752
* /organism="Homo sapiens"
* /db_xref="taxon:9606"
* /chromosome="6"
* /clone_lib="RPC1-11.2"
* 1..3250
* /note="assembly_fragment:01752
* fragment_chain:1"
* 3351..14600
* /note="assembly_fragment:01177
* fragment_chain:1"
* 14701..32357
* /note="assembly_fragment:00673
* fragment_chain:1"
* 32458..34886
* /note="assembly_fragment:00884
* fragment_chain:1"
* 34987..43490
* /note="assembly_fragment:00652
* fragment_chain:1"
* 43591..47437
* /note="assembly_fragment:00628
* fragment_chain:1"
* 47538..57356
* /note="assembly_fragment:00983
* fragment_chain:2"
* 57457..59845

```

/note="assembly_fragment:01129
fragment_chain:2"
59946..63972
/note="assembly_fragment:01207
fragment_chain:3"
64073..82711
/note="assembly_fragment:00301
fragment_chain:3"
82812..111814
/note="assembly_fragment:00470"
111915..120276
/note="assembly_fragment:00783.0"
120377..136660
/note="assembly_fragment:01014"
136761..153913
/note="assembly_fragment:01161"
154014..158659
/note="assembly_fragment:01329"
158760..164235
/note="assembly_fragment:01675"
164336..184996
/note="assembly_fragment:01884"
185097..193752
/note="assembly_fragment:01893"
BASE COUNT 55929 a 40409 c 39996 g 55686 t 1732 others
ORIGIN

Query Match 99.5%; Score 299.4; DB 2; Length 193752;
Best Local Similarity 99.7%; Pred. No. 4.1e-76;
Matches 300; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGCTGTGCGCTCTCCAAATTC 60
Db 12082 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGCTGTGCGCTCTCCAAATTC 12023
QY 61 TGGGAAGGACTTCTCAATCCTAGAGTCTCTACCTTATATTTGAGATGATGAGACAGC 120
Db 12022 TGGGAAGGACTTCTCAATCCTAGAGTCTCTACCTTATATTTGAGATGATGAGACAGC 11963
QY 121 CACAAGTCATGGGTTTAAATTTCTTTTCCCATGCATATGCTCAAAAGGAAGTGTCTATG 180
Db 11962 CACAAGTCATGGGTTTAAATTTCTTTTCCCATGCATATGCTCAAAAGGAAGTGTCTATG 11903
QY 181 GCCCTTGCTTTTATTTAACAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
Db 11902 GCCCTTGCTTTTATTTAACAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 11843
QY 241 AATGTCAGCGCGGCACGGTGGCTCACCCCTGTATCCAGCACCTTTGGAGGCCGAGG 300
Db 11842 AATGTCAGCGCGGCACGGTGGCTCACCCCTGTATCCAGCACCTTTGGAGGCCGAGG 11783
QY 301 C 301
Db 11782 C 11782

RESULT 15
AF301592 4349 bp DNA linear MAM 13-DEC-2000
LOCUS AF301592
DEFINITION Diceros bicornis HFE gene, exons 2 through 6 and partial cds.
ACCESSION AF301592
VERSION AF301592.1 GI:11692702
KEYWORDS black rhinoceros.
SOURCE Diceros bicornis
ORGANISM Diceros bicornis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Perissodactyla; Rhinocerotidae; Diceros.
REFERENCE 1 (bases 1 to 4349)
AUTHORS West,C.J., Worley,M. and Beutler,E.
TITLE Rhinoceros HFE Polymorphisms
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4349)

AUTHORS West,C.J., Worley,M. and Beutler,E.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-2000) Molecular and Experimental Medicine, The
Scripps Research Institute, 10550 North Torrey Pines Road, La
Jolla, CA 92037, USA
FEATURES
source Location/Qualifiers
1..4349 /organism="Diceros bicornis"
/db_xref="taxon:9805"
mRNA join(<1..264,468..743,1588..1863,1995..2108,3109..>3149)
/product="HFE"
exon 1..264
/number=2
CDS join(<1..264,468..743,1588..1863,1995..2108,3109..>3149)
/codon_start=3
/product="HFE"
/protein_id="AAG39940.1"
/db_xref="GI:11692703"
/translation="SHSLRYLEWGCASERDHLPLFEALGYVDDELFLVAYNHESRRAES
RAQWLGEAHSQWLQTLQSLKGDHMFIVDFWTIMDNHNSKESHVLTQVILGCEVQE
DNSTRGFWKYGDGDHLEFCPETLDMRAAESRALITKLEWNEVKNIRAKONRAYLEND
CPEQLQWILLELGRVLGDQVPLVKVTHHVASAVTTLRQALNFYQONITRWLKDORR
PVDVKDAESKDPLSPGDGTQSWAALAVPFGEQRYTCQVEHFLDQPLTATWEPSSL
NTLVGTVSGIAVCVITFFIGILFRILKRQASRGAMGDYVLAEC"
exon 468..743
/number=3
exon 1588..1863
/number=4
exon 1995..2108
/number=5
exon 3109..>3149
/number=6
BASE COUNT 1142 a 924 c 1124 g 1159 t
ORIGIN
Query Match 29.5%; Score 88.8; DB 4; Length 4349;
Best Local Similarity 75.9%; Pred. No. 2e-15; Indels 2; Gaps 1;
Matches 123; Conservative 0; Mismatches 37;
QY 87 GTCTCTACCTTATATTTGAGATGATGAGACAGCCACAGTCATGGTGTATTTCTTTT 146
Db 908 GTCTCCACCTTGTCTATGAGATGATGACACAAACCATGATGATGATTTTATTTT 967
QY 147 CTCCTATGCATATGGCTCAAGGGAAGTGTCTATGGCCCTTGTCTTTTATTTAACCAATAA 206
Db 968 CCCCACACATATGGCCAAAGAAGTGTCTGTGATC--TTCTCTGTGTTTAAACCAAGAA 1025
QY 207 TCTTTTGTATATTTATACCTGTTAAAAATTCAGAAATGTCAA 248
Db 1026 CTCTGTGTCTACTTATACATTTTCAGAAATTCAGAGGTGTCAA 1067
Search completed: June 19, 2002, 10:58:56
Job time: 8334 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:24:43 ; Search time 1700.16 Seconds
(without alignments)
2389.530 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870
Perfect score: 301
Sequence: 1 GGCACGGAATCCCTGGTGG.....GCACTTGGGAGCCGAGGC 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estovl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260.2	86.4	531	12	AQ703007 HS_5443_B
2	65.2	21.7	630	12	AG019823 Homo sapi
3	65.2	21.7	659	12	AG020621 Homo sapi
c	5	65	21.6	422	W44873 zc05d01.r1
	5	64	21.3	651	10 BG388094
6	63.6	21.1	533	12	AQ471624 CITBI-E1-
c	7	63.4	21.1	303	12 AZ756856
c	8	63	20.9	389	9 AN825901
c	9	62.8	20.9	342	9 AI092525
c	10	62.8	20.9	439	12 AQ150673
11	62.8	20.9	553	9	AW020486
12	62.2	20.7	281	9	AI565873
13	62.2	20.7	310	10	BF898122
14	62.2	20.7	400	10	R89296
15	62.2	20.7	408	10	H71349
16	62.2	20.7	414	10	N58729
c	17	62.2	20.7	556	12 AQ489145

18	62.2	20.7	603	9	AF063563	
19	62.2	20.7	603	9	AI110688	
c	20	62.2	662	12	AG061477 Pan trogl	
21	62.2	20.7	1825	11	AF113008 Homo sapi	
c	22	61.4	619	9	BE147774 RCI-HW022	
23	61.2	20.3	466	12	AQ269092 RPC111-74	
c	24	61.2	703	12	AG090274 Pan trogl	
25	61	20.3	645	12	AG095800 Pan trogl	
26	60.8	20.2	369	10	F25183 HSPDI2013 H	
27	60.6	20.1	435	9	AA573540 nfa1B05.s	
28	60.6	20.1	455	9	AI732329 nfa1B05.x	
29	60.6	20.1	511	9	AI732239 nf35f10.x	
30	60.6	20.1	693	12	AG144127 Pan trogl	
31	60.4	20.1	656	12	AG115691 Pan trogl	
32	60.4	20.1	657	12	AG084435 Pan trogl	
33	60.2	20.0	363	12	AQ385523 RPC111-13	
34	60.2	20.0	586	9	AV713396	
35	60	19.9	368	9	AA720754 nw93g11.x	
c	36	60	19.9	482	12	AQ631584 RPC1-11-4
c	37	60	19.9	748	10	BI522662 603175686
c	38	60	19.9	822	10	BI906585 603064226
c	39	60	19.9	1110	10	BG257716 602377406
c	40	60	19.9	1378	11	BC008448 Homo sapi
c	41	59.8	19.9	495	10	BG054544 Toa5h12.x
c	42	59.6	19.8	415	12	AQ228137 HS-2014_B
c	43	59.6	19.8	976	10	BM423178 PLATE3_CO
c	44	59.4	19.7	415	12	AQ210277 HS-3249_B
c	45	59.4	19.7	433	12	AQ475579 CITB1-E1-

ALIGNMENTS

RESULT 1

LOCUS AQ703007 531 bp DNA linear GSS 07-JUL-1999
DEFINITION HS_5443_B1_G09_T7A_RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=1019 Col=17 Row=N, DNA sequence.
ACCESSION AQ703007 GI:5412433
VERSION AQ703007.1
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 531)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1019 row: N column: 17
Seq primer: T7
Class: BAC ends
High quality sequence stop: 531.
Location/Qualifiers
1. .531
/organism="Homo sapiens"


```

RESULT 4
W44873
LOCUS
DEFINITION
    422 bp mRNA linear EST 10-OCT-1996
    zc05d01.r1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone
    IMAGE:321409 5' similar to contains Alu repetitive element,, mRNA
    sequence.
ACCESSION
W44873
VERSION
W44873.1 GI:1328963
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 422)
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE
The WashU-Merck EST Project
JOURNAL
Unpublished (1995)
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 597 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 350.
Location/Qualifiers
    1..422
    /organism="Homo sapiens"
    /db_xref="GDB:1259067"
    /db_xref="taxon:9606"
    /clone="IMAGE:321409"
    /clone_lib="Soares_parathyroid_tumor_NbHPA"
    /tissue_type="parathyroid tumor"
    /dev_stage="adult"
    /lab_host="DH10B (ampicillin resistant)"
    /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia
    ) with a modified polylinker; Site_1: Not I; Site_2: Eco
    RI; 1st strand cDNA was primed with a Not I - oligo(dT)
    primer
    [5'-TGTTTACCATCTGAAGTGGGAGCGCGACCAATTTTTTTTTTTTTTTTTTTT
    TTTT-3'], double-stranded cDNA was size selected, ligated
    to Eco RI adapters (Pharmacia), digested with Not I and
    cloned into the Not I and Eco RI sites of a modified pT7T3
    vector (Pharmacia). Library went through one round of
    normalization to a cot = 5. Library constructed by Bento
    Soares and M.Fatima Bonaído. RNA from sporadic parathyroid
    adenomas was kindly provided by Dr. Stephen Marx, National
    Institute of Diabetes and Digestive and Kidney Diseases,
    NIH."
BASE COUNT 117 a 96 c 93 g 106 t 10 others
ORIGIN

Query Match 21.6%; Score 65; DB 10; Length 422;
Best Local Similarity 79.4%; Pred. No. 2.9e-06;
Matches 77; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 205 AATCTTTGTATATTATACCTGTAATAAATTCAGAAATGTCAGGCGCGGACGGTGGC 264
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 AATAGTTTGCCCTCTAGTCCCTTATTAATTAATGAAGGCGTGTGTCATGGTGGC 290

Qy 265 TCACCCCTGAATCCAGCACTTTGGGAGCCGAGGC 301
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 TCAGCGCTGTAAATCCAGCACTTTGGGAGCCGAGGC 327

RESULT 5
BG388094
LOCUS
DEFINITION
    602413054F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4521732 5',
    mRNA sequence.
ACCESSION
BG388094
VERSION
BG388094.1 GI:13281540
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 651)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10421 row: k column: 13
High quality sequence stop: 633.
Location/Qualifiers
    1..651
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:4521732"
    /clone_lib="NIH_MGC_92"
    /tissue_type="embryonal carcinoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
    Average insert size 2.5 kb. Library enriched for
    full-length clones and constructed by Life Technologies.
    Note: this is a NIH_MGC Library."
BASE COUNT 180 a 160 c 137 g 174 t
ORIGIN

Query Match 21.3%; Score 64; DB 10; Length 651;
Best Local Similarity 70.8%; Pred. No. 5.2e-06;
Matches 85; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

Qy 181 GCCCTGCTTTTATTACCAATAATCTTTTGTATATTATACCTGTAAAAATTGAGA 240
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 GCATTCTGTATATGATACCATCTTGATGGTTTATATAGATGATATATCTAAA 307

Qy 241 AATGTCAGGCGGCGACGGTGGCTCACCCCTGTAATCCAGCACTTTGGGAGCGCGAGG 300
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 TACATATAGCTGGAGCGAGTGGCTCACACCTGTAATCCAGCACTTTGGGAGCGCAAGG 247

RESULT 6
AQ471624
LOCUS
DEFINITION
    AQ471624
    CITBI-E1-2590G21.TR CITBI-E1 Homo sapiens genomic clone 2590G21,
    DNA sequence.
ACCESSION
AQ471624
VERSION
AQ471624.1 GI:4655278
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 533)
AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building

```



```

source
1..389
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1371621"
/clone_lib="NCI_CGAP_GCB1"
/tissue_type="germinal center B cell"
/lab_host="DH10B"
/note="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I - oligo(dT) primer
15'-TGTTACCAATCTGAAGTGGGCGCGCTCATTTTTTTTTTTT-3'
1. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaudo."
BASE COUNT      94 a 79 c 89 g 127 t
ORIGIN

Query Match      20.9%; Score 63; DB 9; Length 389;
Best Local Similarity 61.1%; Pred. No. 9.3e-06;
Matches 102; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 135 TTAATTCCTCTCATCATGCTCAAGGGAAGTGTATGCGCTTGCTTTTA 194
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 274 TTAGTGAGTTTATTCAGACCAAGTGGTGAATCTGAAATGTGATTCACATGATGA 215
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 195 TTAAACCAATATCTTTTGATATTATTAACCTGTGTTAAAAATTCAGAAATGCAAGGCCGG 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 214 TTCAATGACTACCCAAATGAGTTTATAACTGATCAAAAAATTTAAAAATCCAAGCCGG 155
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 GCACGGTGTGCTCACCCTGTAATCCAGCACCTTGGAGGCCGAGGC 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 154 GCATGTGGCTCACACTGTAATCCAGCACCTTGGAGGCCGAGGC 108
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
AI092525/c
LOCUS
DEFINITION
q38f07.x1 Soares_NHMPu_SI Homo sapiens cDNA clone IMAGE:1689061
3' similar to contains Alu repetitive element.; mRNA sequence.
ACCESSION
AI092525
VERSION
AI092525.1 GI:3431501
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 342)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham.
FEATURES
Location/Qualifiers
1..342
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1689061"
/clone_lib="Soares_NHMPu_SI"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: p7T3D-Pac

```

```

(Pharmacia) with a modified polylinker; Site.1: Not I;
Site.2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT      109 a 72 c 75 g 86 t
ORIGIN

Query Match      20.9%; Score 62.8; DB 9; Length 342;
Best Local Similarity 71.9%; Pred. No. 1e-05;
Matches 82; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 188 CTTTTTATTAACCAATAATCTTTTGTATATTATACCTGTATAAAATTCAGAAATGTCA 247
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 CTGTTTATTTTCAGATCTTTCTTCCTTCATCTCCGATAAAAGAGATATGAAA 123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 248 AGCGCGGCGGCTGCTCACCCTGTAATCCAGCACCTTGGAGGCCGAGGC 301
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 AGCGCGGCGCAATGCTCAGCCTGTAATCCAGCACCTTGGAGGCCGAGGC 69
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AI050673/c
LOCUS
DEFINITION
H5_3203_A2_A05_MR CIT Approved Human genomic Sperm Library D Homo
sapiens genomic clone Plate-3203 Col-10 Row=A, DNA sequence.
ACCESSION
AI050673
VERSION
AI050673.1 GI:3544131
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 439)
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3203 row: A column: 10
Class: BAC ends
High quality sequence stop: 439.
FEATURES
Location/Qualifiers
1..439
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="plate-3203 Col-10 Row=A"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      126 a 103 c 95 g 114 t 1 others
ORIGIN

Query Match      20.9%; Score 62.8; DB 12; Length 439;
Best Local Similarity 59.2%; Pred. No. 1e-05;

```



```

RESULT 14
R89296 400 bp mRNA linear EST 24-AUG-1995
LOCUS YQ02fl1.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
DEFINITION IMAGE:195789 3' similar to contains Alu repetitive element; contains
L1 repetitive element ;, mRNA sequence.
R89296
ACCESSION R89296.1 GI:954123
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 400)
AUTHORS Hillier,L., Clark,L., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The Washo-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1115
High quality sequence stops: 268
Source: IMAGE Consortium, LLMNL
This clone is available royalty-free through LLMNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1115 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 268.
location/Qualifiers
1. .400
/organism="Homo sapiens"
/db_xref="GDB:3764839"
/db_xref="taxon:9606"
/clone="IMAGE:195789"
/clone.lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="organ: Liver and Spleen; Vector: pYT3D (Pharmacia
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5', AACCTGGAGAATAAATAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac
and Eco RI sites of the modified pYT3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 119 a 83 c 112 g 81 t
ORIGIN
Query Match 20.7%; Score 62.2; DB 10; Length 400;
Best Local Similarity 76.8%; Pred. No. 1.5e-05;
Matches 76; Conservative 0; Mismatches 23; Indels 0; Gaps 0
QY 203 ATATCTTTTGATATTTTACCTGTTAAAAATTCAGAAATGTC AAGCGCGGCACGCGT 262
Db 82 ACATACATTCAGATATATTTTCACCTTAAAAAGAGAAATCATGCGCGCGCGGTG 141
QY 263 GCTCACCCCTGTAATCCACGACTTTGGAGGCGCGAGGC 301
Db 142 GCTCACGCTGTAATCCACGACTTTGGAGGCGCGAGAC 180

```

Search completed: June 19, 2002, 10:24:39

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:54:02 ; Search time 213.18 Seconds
(without alignments)
2424.201 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870
Perfect score: 301
Sequence: 1 GGCACGGAACTCCCTGGTTGG.....GCACCTTGGAGGCCGAGGC 301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	301	100.0	10825	18	AAT96690 Hereditary haemoch
2	301	100.0	10825	22	AAC68425 Human hereditary h
3	301	100.0	10825	22	AAC68426 Human hereditary h
4	301	100.0	10825	22	AAC68427 Human hereditary h
5	301	100.0	10825	21	AAC68428 Human hereditary h
6	301	100.0	12146	21	AAA96794 Genomic DNA of a h
7	299.4	99.5	5749	22	AAL36747 Human musculoskele
8	294.6	97.9	237326	19	AAV57903 Hereditary haemoch
9	61.4	20.4	429	21	AAH30227 Human colon cancer

10	61.4	20.4	26110	22	AAK65036 Human immune/haema
11	61.4	20.4	26110	22	AAK78526 Human immune/haema
12	61	20.3	584	22	AAK79701 Human immune/haema
13	61	20.3	6534	22	AAK02766 Human reproductiv
14	60.6	20.1	344	21	AAC27863 Human secreted pro
15	60.6	20.1	3242	22	AAI86495 Human polynucleoti
16	60.6	20.1	5768	22	AAK05331 Human reproductiv
17	60.6	20.1	7619	22	AAK44934 cDNA encoding nove
18	60	19.9	655	22	AAK56970 Human immune/haema
19	60	19.9	42519	22	AAK81318 Human immune/haema
20	59.8	19.9	448	22	AAI93368 Human polynucleoti
21	59.6	19.8	841	22	AAK85928 Human immune/haema
22	59.6	19.8	22655	22	AAK70122 Human immune/haema
23	59.2	19.7	16595	20	AAK23521 Human kidney amino
24	59	19.6	606	22	AAK04682 Human reproductiv
25	59	19.6	5270	22	AAK64901 Human immune/haema
26	58.8	19.5	328	21	AAC30053 Human secreted pro
27	58.8	19.5	4156	22	AAK83412 Human immune/haema
28	58.6	19.5	1146	20	AAH85051 Human secreted pro
29	58.6	19.5	4512	22	AAH16662 Human cDNA sequenc
30	58.6	19.5	17979	22	AAK64964 Human immune/haema
31	58.6	19.5	17979	22	AAK71664 Human immune/haema
32	58.4	19.4	270	22	AAK56284 Human immune/haema
33	58.4	19.4	338	22	AAK38872 Novel human diagno
34	58.4	19.4	484	22	AAI93692 Human polynucleoti
35	58.4	19.4	697	22	AAH70948 Human cervical can
36	58.4	19.4	986	21	AAC57483 Arachidonic acid m
37	58.4	19.4	1641	22	ABAI6008 Human nervous syst
38	58.4	19.4	6004	22	AAK84542 Human immune/haema
39	58.4	19.4	14969	22	AAK78763 Human immune/haema
40	58.4	19.4	21837	22	AAK85946 Human 9p11 chromos
41	58.4	19.4	57728	22	AAK87588 Human 9p11 chromos
42	58.2	19.3	710	22	AAH98518 Human EST-derived
43	58.2	19.3	32184	22	AAK05850 Human reproductiv
44	58.2	19.3	32204	22	AAK05849 Human reproductiv
45	58.2	19.3	51474	22	AAK97846 Human neuroblastom

ALIGNMENTS

RESULT 1					
AAT96690					
ID	AAT96690	standard; DNA; 10825 BP.			
XX					
AC	AAT96690;				
XX					
DT	14-APR-1998	(first entry)			
XX					
DE	Hereditary haemochromatosis gene.				
XX					
KW	Hereditary haemochromatosis; metal toxicity; diagnosis;				
KW	gene therapy; prenatal screening; human; ds.				
XX					
OS	Homo sapiens.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	361..7147			
FT		/*tag= a			
FT		/*note= "contains introns"			
FT	intron	437..3761			
FT		/*tag= b			
FT	intron	/*number= 1			
FT		4026..4234			
FT		/*tag= c			
FT		/*number= 2			
FT	intron	4511..5605			
FT		/*tag= d			
FT		/*number= 3			
FT	intron	5882..6039			
FT		/*tag= e			
FT		/*number= 4			
FT	intron	6154..7106			

FT mutation /*tag= f /number= 5 3872

FT /*tag= g /*tag= "C to G substitution (24d2 mutation) results in His to Asp substitution"

FT variation 3878 /*tag= h /*tag= "A to T substitution (24d7 variant) results in Ser to Cys substitution"

FT mutation 5834 /*tag= i /*tag= "G to A substitution (24d1 mutation associated with HH), results in Cys to Tyr substitution"

XX WO9738137-A1.

XX 16-OCT-1997.

XX 04-APR-1997; 97WO-US06254.

XX 23-MAY-1996; 96US-0652265.

PR 04-APR-1996; 96US-0630912.

PR 16-APR-1996; 96US-0632673.

XX (MERC-) MERCATOR GENETICS INC.

XX Drayna DT, Feder JN, Gnirke A, Ruddy D, Thomas WJ; Tsuchihashi Z, Wolff RK;

XX WPI; 1997-512743/47.

DR P-PSDB; AAW36499.

XX Hereditary haemochromatosis gene and variants - useful for diagnosis and treatment of hereditary haemochromatosis disease

XX Disclosure; Fig 3; 115pp; English.

CC This genomic DNA sequence corresponds to the human gene whose mutated form is associated with hereditary haemochromatosis (HH). To identify this novel gene, allelic association patterns were determined between known markers and the HH locus in the HLA region of chromosome 6. A physical clone coverage was then generated extending from D6S265, which is a marker that is centromeric of HNA-A, in a telomeric direction through D6S276, a marker at which the allelic association was no longer observed. A single mutation (24d1) in the HH gene appears responsible for the majority of HH disease. This comprises a G to A substitution that is present in 88% of affected chromosomes and in 4% of unaffected chromosomes. It results in a Cys to Tyr substitution in the encoded protein (see AAW36499) at a critical disulphide bridge important for secondary structure. The following are claimed: the HH genomic DNA (1), a 1437 bp cDNA sequence (1a) (see AAT96691) and their 24d1, 24d2 and 24d7 variants; a cloning or expression vector; host cells; a peptide product chosen from the HH gene product, its variants (24d1, 24d2 and 24d7), or a peptide of at least 56 amino acid residues of these; an antibody produced using the peptide; a method to determine the presence or absence of the common HH gene mutation; an animal model for the HH disease; metal chelation agents, T-cell differentiation factors and therapeutic agents for the mitigation of injury due to oxidative process in vivo or mitigation of iron overload; a method for screening potential therapeutic agents for activity in connection with HH disease; an antisense oligonucleotide directed against a transcriptional product of a nucleic acid sequence as above; and oligonucleotides or pairs of oligonucleotides covering a range of nucleotides from (1), (1a) or their variants, useful for detecting a polymorphism in the HH gene. The invention also relates to methods for screening for HH homozygotes, to HH diagnosis, prenatal screening and diagnosis, and therapies of HH disease, including gene therapy, protein- and antibody-based therapeutics, and small molecule therapeutics.

XX SQ Sequence 10825 BP; 2996 A; 2254 C; 2648 G; 2927 T; 0 other;

Query Match 100.0%; Score 301; DB 18; Length 10825; Best Local Similarity 100.0%; Pred. No. 1e-79; Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGGAATCCTGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
|||||
Db 4570 ggcacggaatccctggttgagattccagagtggtgagctgtgtgctctccaaattc 4629

QY 61 TGGGAAGGAGCTTTCTCAATCTAGAGTCTCTACTTAAATGAGATGATGAGACAGC 120
|||||
Db 4630 tgggaaggagactttctcaatcctagagtcctacacctataattgagatgtatgagacgc 4689

QY 121 CACAAGTCATGGGTTTAATTTCTTCTCCATGATATGGCTCAAGGAAAGTGTCTATG 180
|||||
Db 4690 cacaagtcatgggttctaattctctcccatgcatacggctcaaaaggaggtgtccatg 4749

QY 181 GCCCTTGCTTTTATTAAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
|||||
Db 4750 gcccttgctttttatttaaccaataatctttgtatatttatactgttaaaattccaga 4809

QY 241 AATGTCAAGGCGGCGACGGTGGCTCACCCCTGTAAATCCAGCACACTTTGGAGGCCGAGG 300
|||||
Db 4810 aatgtcaaggcgggcacggtggctcaccctgtaatccacagcactttgggagggccgagg 4869

QY 301 C 301
Db 4870 c 4870

RESULT 2
AAC68425
ID AAC68425 standard; DNA; 10825 BP.

XX AC AAC68425;

XX DT 21-FEB-2001 (first entry)

XX DE Human hereditary hemochromatosis DNA.

XX KW HH; hereditary hemochromatosis; chelation agent;

XX KW T-cell differentiation factor; iron overload; ds.

XX OS Homo sapiens.

XX PN US6140305-A.

XX PD 31-OCT-2000.

XX PF 04-APR-1997; 97US-0834497.

XX PR 04-APR-1996; 96US-0630912.

XX PR 16-APR-1996; 96US-0632673.

XX PR 23-MAY-1996; 96US-0652265.

XX PA (BIRA) BIO-RAD LAB INC.

XX PI Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;

XX PI Feder JN;

XX DR WPI; 2001-006341/01.

XX DR P-PSDB; AAW36869.

XX PT New hereditary hemochromatosis gene products or polypeptides, useful for treating hereditary hemochromatosis in a patient, and as a metal chelation agent alleviating iron overload -

XX PS Disclosure; Fig 3; 108pp; English.

XX CC The present invention relates to hereditary hemochromatosis gene

CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2998 A; 2253 C; 2648 G; 2926 T; 0 other;

Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGGCTGAGGCTGTGCCTCTCCAAATTC 60
Db 4570 ggcacggaatccctgggttgagtttcagagtggtgaggtgtgctctccaaattc 4629

QY 61 TGGGAAGGACTTTCGAATCCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
Db 4630 tgggaaggactttctcaatccctagagttctctaccctataattgagatgtatgagacagc 4689

QY 121 CACAAGTCATGGGTTTAATTTCTTTCTCATGCAATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 cacaagtcattgggtttaaattctttctccatgcataatggctcaagggaagtgtctatg 4749

QY 181 GCCCTTGCTTTTATTTAACCATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 240
Db 4750 gcccttgcttttatttaaccataatctttgtatattataacctgttaaaattcaga 4809

QY 241 AATGTCGAAGCGCGGCACGGTGGCTCACCCCTGTAATCCAGCACACTTTGGGAGCGCAGG 300
Db 4810 aatgtcaagcgggcacggttgctcaccctgtaatccagcactttggagcgaggg 4869

QY 301 C 301
Db 4870 c 4870

RESULT 3
AAC68426
ID AAC68426 standard; DNA; 10825 BP.
AC AAC68426;
XX
XX
XX 21-FEB-2001 (first entry)
XX Human hereditary hemochromatosis 24d1 mutation DNA.
XX HH; hereditary hemochromatosis; chelation agent;
XX T-cell differentiation factor; iron overload; ds.
XX
XX Homo sapiens.
XX
XX US6140305-A.
XX
XX 31-OCT-2000.
XX
XX 04-APR-1997; 97US-0834497.
XX
XX 04-APR-1996; 96US-0630912.
XX
XX 16-APR-1996; 96US-0632673.
XX
XX 23-MAY-1996; 96US-0652265.
XX
XX (BIRA) BIO-RAD LAB INC.
XX
XX Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
XX Feder JN;
XX
XX WPI; 2001-006341/01.
XX
XX P-PSDB; AAB36870.
XX
XX New hereditary hemochromatosis gene products or polypeptides, useful
XX for treating hereditary hemochromatosis in a patient, and as a metal

PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 3; 108pp; English.
XX
XX The present invention relates to hereditary hemochromatosis gene
XX products. These proteins may be used to treat a patient diagnosed as
XX having human hemochromatosis disease. It is also useful as a metal
XX chelation agent or as a T-cell differentiation factor, and for
XX alleviating iron overload. They may also be used in protein replacement
XX therapy for individuals having a defective human hemochromatosis gene.
XX
SQ Sequence 10825 BP; 2999 A; 2253 C; 2647 G; 2926 T; 0 other;

Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGGCTGAGGCTGTGCCTCTCCAAATTC 60
Db 4570 ggcacggaatccctgggttgagtttcagagtggtgaggtgtgctctccaaattc 4629

QY 61 TGGGAAGGACTTTCGAATCCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
Db 4630 tgggaaggactttctcaatccctagagttctctaccctataattgagatgtatgagacagc 4689

QY 121 CACAAGTCATGGGTTTAATTTCTTTCTCATGCAATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 cacaagtcattgggtttaaattctttctccatgcataatggctcaagggaagtgtctatg 4749

QY 181 GCCCTTGCTTTTATTTAACCATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 240
Db 4750 gcccttgcttttatttaaccataatctttgtatattataacctgttaaaattcaga 4809

QY 241 AATGTCGAAGCGCGGCACGGTGGCTCACCCCTGTAATCCAGCACACTTTGGGAGCGCAGG 300
Db 4810 aatgtcaagcgggcacggttgctcaccctgtaatccagcactttggagcgaggg 4869

QY 301 C 301
Db 4870 c 4870

RESULT 4
AAC68427
ID AAC68427 standard; DNA; 10825 BP.
XX
XX AAC68427;
XX
XX 21-FEB-2001 (first entry)
XX
XX Human hereditary hemochromatosis 24d2 mutation DNA.
XX HH; hereditary hemochromatosis; chelation agent;
XX T-cell differentiation factor; iron overload; ds.
XX
XX Homo sapiens.
XX
XX US6140305-A.
XX
XX 31-OCT-2000.
XX
XX 04-APR-1997; 97US-0834497.
XX
XX 04-APR-1996; 96US-0630912.
XX
XX 16-APR-1996; 96US-0632673.
XX
XX 23-MAY-1996; 96US-0652265.
XX
XX (BIRA) BIO-RAD LAB INC.
XX
XX Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
XX Feder JN;
XX

```
DR WPI: 2001-006341/01.
DR P-PSDB: AAB36871.
XX
XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 3; 108pp; English.
XX
XX The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
XX Sequence 10825 BP; 2998 A; 2252 C; 2649 G; 2926 T; 0 other;
SQ
Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCTCTCCAAATTC 60
DB 4570 ggcacggaatccctggttgagtttcagaggtggctgaggtgtgtgctctccaaattc 4629
QY 61 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTTAATTTAGATGTATGAGACAGC 120
DB 4630 tgggaaggagcttctcaatcctagagtcctaccttataattgagatgtatgagacagc 4689
QY 121 CACAGTCATGGGTTAATTTCTCCATGCATATGGCTCAAGGAGAGTGTCTATG 180
DB 4690 cacaagtcattgggttaatttctccatgcataatgctcaaaagggaagtgtctatg 4749
QY 181 GCCCTTGCTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
DB 4750 gcccttgcttttatttaaccaataatctttgtatattatactgttaaaattcaga 4809
QY 241 AATGTCAGCGCGGACGGTGGCTCACCCCTGTAATCCAGCACTTTGGGAGCGCGAGG 300
DB 4810 aatgtcaaggcggcggtggctcacccctgtaatccagcactttgggagggcggagg 4869
QY 301 C 301
DB 4870 c 4870
RESULT 5
AAC68428
ID AAC68428 standard; DNA; 10825 BP.
XX
XX AAC68428;
XX
XX 21-FEB-2001 (first entry)
XX
XX Human hereditary hemochromatosis 24d1/2 mutation DNA.
XX
XX HH; hereditary hemochromatosis; chelation agent;
KW T-cell differentiation factor; iron overload; ds.
XX
XX Homo sapiens.
OS
XX
XX US6140305-A.
PN
XX
XX 31-OCT-2000.
PD
XX
XX 04-APR-1997; 97US-0834497.
PF
XX
XX 04-APR-1996; 96US-0630912.
PR
XX
XX 26-APR-1996; 96US-0632673.
PR
XX
XX 23-MAY-1996; 96US-0652265.
FT
XX
```

```
PA (BIRA ) BIO-RAD LAB INC.
XX
XX Thomas WJ, Drayna DT, Gnirke A, Ruddy D, Tsuchihashi Z, Wolff RK;
PI Feder JN;
XX
XX WPI: 2001-006341/01.
DR P-PSDB: AAB36872.
XX
XX New hereditary hemochromatosis gene products or polypeptides, useful
PT for treating hereditary hemochromatosis in a patient, and as a metal
PT chelation agent alleviating iron overload -
XX
XX Disclosure; Fig 3; 108pp; English.
XX
XX The present invention relates to hereditary hemochromatosis gene
CC products. These proteins may be used to treat a patient diagnosed as
CC having human hemochromatosis disease. It is also useful as a metal
CC chelation agent or as a T-cell differentiation factor, and for
CC alleviating iron overload. They may also be used in protein replacement
CC therapy for individuals having a defective human hemochromatosis gene.
XX
XX Sequence 10825 BP; 2999 A; 2252 C; 2648 G; 2926 T; 0 other;
SQ
Query Match 100.0%; Score 301; DB 22; Length 10825;
Best Local Similarity 100.0%; Pred. No. 1e-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCTCTCCAAATTC 60
DB 4570 ggcacggaatccctggttgagtttcagaggtggctgaggtgtgtgctctccaaattc 4629
QY 61 TGGGAAGGAGCTTCTCAATCCTAGAGTCTCTACCTTTAATTTAGATGTATGAGACAGC 120
DB 4630 tgggaaggagcttctcaatcctagagtcctaccttataattgagatgtatgagacagc 4689
QY 121 CACAGTCATGGGTTAATTTCTCCATGCATATGGCTCAAGGAGAGTGTCTATG 180
DB 4690 cacaagtcattgggttaatttctccatgcataatgctcaaaagggaagtgtctatg 4749
QY 181 GCCCTTGCTTTTATTTAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
DB 4750 gcccttgcttttatttaaccaataatctttgtatattatactgttaaaattcaga 4809
QY 241 AATGTCAGCGCGGACGGTGGCTCACCCCTGTAATCCAGCACTTTGGGAGCGCGAGG 300
DB 4810 aatgtcaaggcggcggtggctcacccctgtaatccagcactttgggagggcggagg 4869
QY 301 C 301
DB 4870 c 4870
RESULT 6
AAA96794
ID AAA96794 standard; cDNA; 12146 BP.
XX
XX AAA96794;
XX
XX 19-FEB-2001 (first entry)
XX
XX Genomic DNA of a histocompatibility iron loading (HFE) gene.
XX
XX Human; histocompatibility iron loading protein; HFE protein;
KW major histocompatibility complex; non-classical class I gene;
KW chromosome 6p; iron disorder; haemochromatosis; ss.
XX
XX Homo sapiens.
OS
XX
XX Key
FH 1028..1324
FT exon
FT /*tag= a
FT /number= 1
FT
```


FT intron 1325..4651
FT /*tag= b
FT /number= 1
FT 4652..4915
FT /*tag= c
FT /number= 2
FT 4916..5124
FT /*tag= d
FT /number= 2
FT 5125..5400
FT /*tag= e
FT /number= 3
FT 5401..6493
FT /*tag= f
FT /number= 3
FT 6494..6769
FT /*tag= g
FT /number= 4
FT 6770..6927
FT /*tag= h
FT /number= 4
FT 6928..7041
FT /*tag= i
FT /number= 5
FT 7042..7994
FT /*tag= j
FT /number= 5
FT 7995..9050
FT /*tag= k
FT /number= 6
FT 9051..10205
FT /*tag= l
FT /number= 6
FT 10206..10637
FT /*tag= m
XX WO200058515-A1.
XX
XX 05-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US07982.
XX
XX 26-MAR-1999; 99US-0277457.
XX (BILL-) BILLUPS-ROTHENBERG INC.
XX
XX Rothenberg BE, Sawada-Hirai R, Barton JC;
XX WPI; 2000-647244/62.
XX
XX Diagnosing an iron disorder e.g. hemochromatosis or a genetic
XX susceptibility to develop it, by determining the presence of a mutation
XX in exon 2 or an intron of a histocompatibility iron loading nucleic
XX acid -
XX
XX Example 1; Page 21-28; 55pp; English.
XX
XX The present sequence represents the human histocompatibility iron
XX loading (HFE) gene. The HFE gene is a major histocompatibility (MHC)
XX non-classical class I gene located on chromosome 6p. Mutations in the
XX gene lead to iron disorders. The specification describes a method for
XX diagnosing an iron disorder or a genetic susceptibility to develop the
XX disorder in a mammal. The method comprises determining the presence of
XX a mutation in exon 2 or an intron of a HFE gene or protein. The mutation
XX is not a C to G missense mutation at nucleotide 187 of the sequence
XX given in A96769 (Genbank Accession number U60319). The presence of the
XX mutation indicates the disorder or the genetic susceptibility to the
XX disorder. The method is used to diagnose an iron disorder
XX e.g. haemochromatosis, or a genetic susceptibility to develop it.

Sequence 12146 BP; 3383 A; 2474 C; 2911 G; 3378 T; 0 other;

Query Match 100.0%; Score 301; DB 21; Length 12146;
Best Local Similarity 100.0%; Pred. No. 1.le-79;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAATCCCTGGTTGGAGCTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 5460 ggcacggaatccctgggtggagtttcagaggtggctgaggtgtgtgctctccaaattc 5519
QY 61 TGGGAAGGGACTTTCGAATCCTAGAGTCTCTACCTTATAATTGAGATGCTATGAGACAGC 120
Db 5520 tgggaagggaactttcgaatcctagagtcctaccttaataattgagatgtatgagacagc 5579
QY 121 CACAGTCATGGGTTAAATTTCTTCCATGCATATGGCTCAAGGAAGTGTCTATG 180
Db 5580 cacaagtcattgggttaattcttctccatgcataatggtcacaagggaagtgtctatg 5639
QY 181 GGCCTTGCTTTTATTAAACCAATAATCTTTTGTTGTTATTTATACCTGTAAAAATTCAGA 240
Db 5640 gcccttgcttttatttaaccaataatcttctgtatatttaacctgttaaaaaattcaga 5699
QY 241 AATGTCGAAGCGCGGACGGTGGCTCACCCCTGTAATCCAGCACTTTGGGAGCGCGAGG 300
Db 5700 aatgtcaaggccgggcacggtggctcacccctgtaatccagcactttgggagggccgagg 5759
QY 301 C 301
Db 5760 c 5760
RESULT 7
AAL36747
ID AAL36747 standard; DNA; 5749 BP.
XX
AC AAL36747;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3112.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234977.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-451937/48.
XX
XX Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the musculoskeletal system including
PT musculoskeletal cancers and also for testing and detection e.g.
PT diagnosis -
XX
PS Example 2; SEQ ID NO 3112; 781pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AAL34669-AAL37666) and proteins
CC (AB03087-AB04109) associated with the musculoskeletal system useful
CC for preventing, treating or ameliorating medical conditions e.g. by
CC protein or gene therapy. The genes are isolated from a range of human
CC tissues disclosed in the specification. The nucleic acids, proteins,
CC antibodies and (ant)agonists are useful in the diagnosis, treatment
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and
CC other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and

CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5749 BP; 1600 A; 1192 C; 1403 G; 1553 T; 1 other;

Claim 1; Fig 9; 209pp; English.

The present invention describes hereditary haemochromatosis gene products from the human haemochromatosis gene. The present sequence represents a hereditary haemochromatosis subregion from an hereditary haemochromatosis (HH) affected individual. Also described is a method to determine the presence or absence of the common hereditary haemochromatosis (HFE) gene mutation in an individual comprising: (a) providing DNA or RNA from the individual; and (b) assessing the DNA or RNA for the presence or absence of a haplotype or genotype where the presence or absence of the haplotype genotype indicates the likely presence of the HFE gene mutation in the genome of the individual. The HFE gene sequences from the present invention can be used to develop products for use in the diagnosis and treatment of HFE. The present invention also describes BfR genes, which are homologues of the milk protein butyrophilin (BT), and can be used in the production of agonists and antagonists of BT function. Also described are: (1) a RoRet gene which can be used to develop products for the study, diagnosis and treatment of lupus and Sjoren's syndrome; and (2) NPT3 and NPT4 genes which are homologues of a type 1 sodium transport gene, and can similarly be used for hypophosphatemia.

Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0215647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX Disclosure; SEQ ID NO 33338; 3071pp + Sequence Listing; English.
PS AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 26110 BP; 7464 A; 4942 C; 5567 G; 8137 T; 0 other;

Query Match 20.4%; Score 61.4; DB 22; Length 26110;
Best Local Similarity 72.1%; Pred. No. 5.3e-08;
Matches 80; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 190 TTTTATTACCAATAATCTTTTGTATATATATACCTGTTAAATAATTCAGAAATGTCAAG 249
Db 8013 tcgtttttatccaactttttaatgtttcttcaaatcctcttaaaactaagaatttttagg 8072
QY 250 GCGGGCAGCGTGCTCACCCGTGTAATCCAGCAGCTTTGGAGGCCGAGS 300
Db 8073 gctggacacggtggtccacgctgtaatccacgactctgagggtcgagg 8123

RESULT 12
AAK79701
ID AAK79701 standard; DNA; 584 BP.
XX
AC AAK79701;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34513.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180828.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.


```
XX AAC27863;
XX AC
XX XX
XX 06-OCT-2000 (first entry)
XX DE
XX Human secreted protein 5' EST, SEQ ID NO: 31938.
XX KW
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.
XX OS
XX Homo sapiens.
XX PN
XX EP1033401-A2.
XX PD
XX 06-SEP-2000.
XX PF
XX 21-FEB-2000; 2000EP-0200610.
XX PR
XX 26-FEB-1999; 9905-0122487.
XX PA
XX (GEST ) GENSET.
XX PI
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX DR
XX WPI; 2000-500381/45.
XX PT
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX PS
XX Claim 1; SEQ ID 31938; 71pp + CD-ROM; English.
XX CC
XX The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX CC identified within the present sequence. The 5' ESTs were prepared from
XX CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX CC sequences usually correspond mainly to the 3' untranslated region (UTR)
XX CC of the mRNA because they are often obtained from oligo-dT primed cDNA
XX CC libraries. Such ESTs are not well suited for isolating cDNA sequences
XX CC derived from the 5' ends of mRNAs and even in those cases where longer
XX CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX CC They are used to obtain upstream regulatory sequences and to design
XX CC expression and secretion vectors.
XX SQ
XX Sequence 344 BP; 96 A; 79 C; 83 G; 85 T; 1 other;

Query Match 20.1%; Score 60.6; DB 21; Length 344;
Best Local Similarity 70.4%; Pred. No. 1.8e-08;
Matches 81; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 186 TGTCTTTTATTTAACCAATAATCTTTTGTATATTTATCTGTATAAAATTCAGAAATCT 245
Db 137 ttcctgcagttccaccacattttcttaattggaataattgggaagaaagaatca 196
QY 246 CAAAGCCGGCAGCGTGGCTACCCCTGTAAATCCAGCAGCTTTGGGAGCCGAGG 300
Db 197 taagccgggcacggtggtcatatctgtatccccagcactttggaggctgag 251

RESULT 15
AA186495/C
ID AA186495 standard; cDNA; 3242 BP.
XX AC
XX AA186495;
XX 06-NOV-2001 (first entry)
XX DT
XX Human polynucleotide SEQ ID NO 6555.
XX DE
XX
```

```
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200164835-A2.
XX PD
XX 07-SEP-2001.
XX PF
XX 26-FEB-2001; 2001WO-US04927.
XX PR
XX 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PA
XX (HYSE-) HYSEQ INC.
XX PI
XX Tang YT, Liu C, Drmanac RT;
XX DR
XX WPI: 2001-514838/56.
XX DR P-PSDB; AAC06564.
XX PT
XX Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders -
XX PS
XX Claim 1; SEQ ID NO 6555; 1399pp + Sequence Listing; English.
XX CC
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ
XX Sequence 3242 BP; 949 A; 656 C; 666 G; 971 T; 0 other;
```

```
Query Match 20.1%; Score 60.6; DB 22; Length 3242;
Best Local Similarity 79.1%; Pred. No. 4.2e-08;
Matches 72; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 211 TTGTATATTTATCTGTATAAAATTCAGAAATGTCAAGCGCGGCGGTGGCTCACC 270
Db 328 TTGTATCTCTACTGCACTACAAATAAGATGTGCGAGGCTGGGCAGGCTCAGC 269
QY 271 CTGTAATCCAGCAGCTTTGGGAGCCGAGGC 301
Db 268 CTGTAATCCAGCAGCTTTGGGAGCCGAGGC 238

Search completed: June 19, 2002, 11:01:22
Job time: 7640 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2002, 08:41:02 ; Search time 46.2 Seconds
(without alignments)
1600.339 Million cell updates/sec

Title: US-09-497-957-3_COPY_4570_4870

Perfect score: 301

Sequence: 1 GCACGGATCCCTGGTGG.....GCACTTGGGAGCGGAGCG 301

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	301	100.0	10825	3	US-08-652-265-1
2	301	100.0	10825	3	US-08-652-265-3
3	301	100.0	10825	3	US-08-652-265-5
4	301	100.0	10825	3	US-08-652-265-7
5	301	100.0	10825	3	US-08-834-497A-1
6	301	100.0	10825	3	US-08-834-497A-3
7	301	100.0	10825	3	US-08-834-497A-5
8	301	100.0	10825	3	US-08-834-497A-7
9	301	100.0	10825	4	US-09-503-444A-1
10	301	100.0	10825	4	US-09-503-444A-3
11	301	100.0	10825	4	US-09-503-444A-5
12	301	100.0	10825	4	US-09-503-444A-7
13	301	100.0	12146	4	US-09-277-457-27
14	301	100.0	246240	2	US-08-724-394A-20
15	301	100.0	246240	2	US-08-724-394A-21
16	301	100.0	246240	2	US-08-724-394A-22
17	58	19.3	87350	3	US-08-781-891-79
18	57.2	19.0	56516	2	US-08-996-306-1
19	57.2	19.0	56516	4	US-09-338-907-1
20	57.2	19.0	56516	4	US-09-218-207-1
21	57.2	19.0	56520	4	US-09-338-907-179
22	57.2	19.0	56520	4	US-09-218-207-179
23	55.4	18.4	7210	2	US-08-257-963B-10
24	55.4	18.4	7210	4	US-08-367-841A-10
25	55.4	18.4	7210	5	PCT-US95-07201-10
26	55.4	18.4	22481	4	US-08-367-841A-43
27	55.4	18.4	22481	5	PCT-US95-07201-43

28	55.2	18.3	780	4	US-09-385-982-32	Sequence 32, Appli
29	55	18.3	2808	3	US-08-870-126-7	Sequence 7, Appli
30	54.8	18.2	14796	4	US-08-975-080-35	Sequence 35, Appli
31	54.8	18.2	14796	4	US-09-630-706-10	Sequence 10, Appli
32	54.8	18.2	14796	4	US-09-496-694B-3	Sequence 3, Appli
33	54.2	18.0	152331	3	US-09-128-155-16	Sequence 16, Appli
34	54.2	18.0	176373	3	US-09-128-155-17	Sequence 17, Appli
35	54	17.9	3867	4	US-09-347-114A-81	Sequence 81, Appli
36	54	17.9	9734	4	US-09-347-114A-80	Sequence 80, Appli
37	53.8	17.9	8409	4	US-09-167-681-37	Sequence 37, Appli
38	53.6	17.8	1252	4	US-09-305-384-7	Sequence 7, Appli
39	53.6	17.8	1947	2	US-08-989-925-2	Sequence 2, Appli
40	53.6	17.8	2555	2	US-08-960-022-15	Sequence 15, Appli
41	53.6	17.8	6235	4	US-09-305-384-5	Sequence 5, Appli
42	53.6	17.8	6679	4	US-09-305-384-1	Sequence 1, Appli
43	53.4	17.7	2784	1	US-08-471-454-1	Sequence 1, Appli
44	53.4	17.7	2784	2	US-08-466-974-1	Sequence 1, Appli
45	53.4	17.7	2784	2	US-08-471-453-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-652-265-1
; Sequence 1, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
; OTHER INFORMATION:
;

```

; OTHER INFORMATION: /note= "No. 6025130mal or wild-type (unaffected)
; OTHER INFORMATION: Hereditary Hemochromatosis (HH) gene
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) allele
; OTHER INFORMATION: CDNA (SEQ ID NO:9)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d2(C)
; OTHER INFORMATION: allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: normal or wild-type (unaffected) genomic
; OTHER INFORMATION: sequence surrounding variant for 24d1(G)
; OTHER INFORMATION: allele (SEQ ID NO:20)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3872, "c")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d2
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(3878, "a")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d7
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "g")
; OTHER INFORMATION: /phenotype= "normal or wild-type
; OTHER INFORMATION: (unaffected)"
; OTHER INFORMATION: /label= 24d1
; US-08-652-265-1

Query Match 100.0%; Score 301; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. NO. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTTCAGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTCAGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629

QY 61 TGGGAAGGGACTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
Db 4630 TGGGAAGGGACTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689

QY 121 CACAAGTCATGGGTTTAAATTTCTTTCTCCATGCATATGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTTAAATTTCTTTCTCCATGCATATGCTCAAGGGAAGTGTCTATG 4749

QY 181 GCCTTGCCTTTTATTTACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCTTGCCTTTTATTTACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809

QY 241 AATGTCGAAGCCGGGCACGGTGGCTCACCCCTGTAATCCAGCACCTTTGGGAGCCGAGG 300
Db 4810 AATGTCGAAGCCGGGCACGGTGGCTCACCCCTGTAATCCAGCACCTTTGGGAGCCGAGG 4869

QY 301 C 301
Db 4870 C 4870
```

```

RESULT 2
US-08-652-265-3
; Sequence 3, Application US/08652265
; Patent No. 6025130
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 17957-000500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION: mutation"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; FEATURE:
; NAME/KEY: -
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: 24d1 allele cDNA (SEQ ID NO:10)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY: -
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "a")
```

OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d1
US-08-652-265-3

Query Match 100.0%; Score 301; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGCTGAGGCTGTGCTCTCCAAATTC 60
DB 4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGCTGAGGCTGTGCTCTCCAAATTC 4629
QY 61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATATTAATTGAGATGTATGAGACAGC 120
DB 4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATATTAATTGAGATGTATGAGACAGC 4689
QY 121 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAGGAAGTGTCTATG 180
DB 4690 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAGGAAGTGTCTATG 4749
QY 181 GCCCTTGCTTTTATTAACCAATAATCTTTTGTATATTTTACCTGTAAAAATTCAGA 240
DB 4750 GCCCTTGCTTTTATTAACCAATAATCTTTTGTATATTTTACCTGTAAAAATTCAGA 4809
QY 241 AATGTCAGGCGGCGGACGGTGGCTCAACCTGTAAATCCAGCACATTTGGGAGGCCGAGG 300
DB 4810 AATGTCAGGCGGCGGACGGTGGCTCAACCTGTAAATCCAGCACATTTGGGAGGCCGAGG 4869
QY 301 C 301
DB 4870 C 4870

RESULT 3

US-08-652-265-5
Sequence 5, Application US/08652265
Patent No. 6025130
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,265
FILING DATE: 23-MAY-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 17957-000500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION: mutation"
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)"
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: 24d2 allele cDNA (SEQ ID NO:11)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
FEATURE:
NAME/KEY: -
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(G) allele (SEQ ID NO:20)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d2
US-08-652-265-5

Query Match 100.0%; Score 301; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGCTGAGGCTGTGCTCTCCAAATTC 60
DB 4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGCTGAGGCTGTGCTCTCCAAATTC 4629
QY 61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATATTAATTGAGATGTATGAGACAGC 120
DB 4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATATTAATTGAGATGTATGAGACAGC 4689
QY 121 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAGGAAGTGTCTATG 180
DB 4690 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAGGAAGTGTCTATG 4749
QY 181 GCCCTTGCTTTTATTAACCAATAATCTTTTGTATATTTTACCTGTAAAAATTCAGA 240
DB 4750 GCCCTTGCTTTTATTAACCAATAATCTTTTGTATATTTTACCTGTAAAAATTCAGA 4809
QY 241 AATGTCAGGCGGCGGACGGTGGCTCAACCTGTAAATCCAGCACATTTGGGAGGCCGAGG 300
DB 4810 AATGTCAGGCGGCGGACGGTGGCTCAACCTGTAAATCCAGCACATTTGGGAGGCCGAGG 4869
QY 301 C 301
DB 4870 C 4870

RESULT 4
US-08-652-265-7
Sequence 7, Application US/08652265

Patent No. 6025130
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
City: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,265
FILING DATE: 23-MAY-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 17957-000500
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: and 24d2 mutations"
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
OTHER INFORMATION: gene containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles"
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: cDNA containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles
OTHER INFORMATION: (SEQ ID NO:12)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
FEATURE:
NAME/KEY: -
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "g")

OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d1
US-08-652-265-7
Query Match 100.0%; Score 301; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCAGGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
DB 4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629
QY 61 TGGGAAGGGACTTCTCAATCCTAGAGTCTCTACCTTATTAATTGAGATGTATGAGACAGC 120
DB 4630 TGGGAAGGGACTTCTCAATCCTAGAGTCTCTACCTTATTAATTGAGATGTATGAGACAGC 4689
QY 121 CACAAGTCATGGGTTTAATTTCTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
DB 4690 CACAAGTCATGGGTTTAATTTCTTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749
QY 181 GCCCTTGCTTTTATTACCAATAATCTTTTGATATTTATACCTGTATAAAATTCAGA 240
DB 4750 GCCCTTGCTTTTATTACCAATAATCTTTTGATATTTATACCTGTATAAAATTCAGA 4809
QY 241 AATGTCAGGCGGCGACGGTGGCTGCCTCACCCCTGTAATCCAGCACTTTGGGAGGCCGAGG 300
DB 4810 AATGTCAGGCGGCGACGGTGGCTGCCTCACCCCTGTAATCCAGCACTTTGGGAGGCCGAGG 4869
QY 301 C 301
DB 4870 C 4870
RESULT 5
US-08-834-497A-1
Sequence 1, Application US/08834497A
Patent No. 6140305
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/834,497A
FILING DATE: 04-APR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/652,265


```
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION: mutation"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; OTHER INFORMATION: gene 24d1 allele"
; FEATURE:
; NAME/KEY:
; LOCATION: 140..7319
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: 24d1 allele cDNA (SEQ ID NO:10)"
; FEATURE:
; NAME/KEY:
; LOCATION: 3852..3891
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
; FEATURE:
; NAME/KEY:
; LOCATION: 5507..6023
; OTHER INFORMATION: /note= "start and stop positions for
; OTHER INFORMATION: genomic sequence surrounding variant
; OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
; FEATURE:
; NAME/KEY: allele
; LOCATION: replace(5834, "a")
; OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
; OTHER INFORMATION:
; OTHER INFORMATION: /label= 24d1
;
US-08-834-497A-3
;
Query Match 100.0%; Score 301; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAATCCCTGGTGTGAGTTTCAGAGGTGGCTGAGGCTGTGCGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTGTGAGTTTCAGAGGTGGCTGAGGCTGTGCGCTCTCCAAATTC 4629
QY 61 TGGGAAGGAGCTTTCATCTCAGAGTCTCTACCTTATATAGATGATGAGACAGC 120
Db 4630 TGGGAAGGAGCTTTCATCTCAGAGTCTCTACCTTATATAGATGATGAGACAGC 4689
QY 121 CACAAGTCATGGTTTAAATTTCTTCTCCATGCATATGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGTTTAAATTTCTTCTCCATGCATATGCTCAAGGGAAGTGTCTATG 4749
QY 181 GCCCTTGCCTTTTATTTAAACCAATATCTTTTGTATTTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCCTTGCCTTTTATTTAAACCAATATCTTTTGTATTTATACCTGTTAAAAATTCAGA 4809
QY 241 AATGTCAAGCGCGGACCGTGGCTACCCCTGTAATCCACGACTTTGGGAGGCCGAGG 300
Db 4810 AATGTCAAGCGCGGACCGTGGCTACCCCTGTAATCCACGACTTTGGGAGGCCGAGG 4869
```

QY 301 C 301
Db 4870 C 4870

RESULT 7
US-08-834-497A-5
; Sequence 5, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gairke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,497A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/652,265
; FILING DATE: 23-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/630,912
; FILING DATE: 04-APR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0056-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
; LOCATION: 6040..6153, 7107..7147)
; OTHER INFORMATION: /product= "Hereditary Hemochromatosis
; OTHER INFORMATION: mutation"
; OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
; OTHER INFORMATION: gene 24d2 allele"
; FEATURE:


```
/ NAME/KEY: -
/ LOCATION: 140..7319
/ OTHER INFORMATION: /note= "start and stop positions for
/ OTHER INFORMATION: 24d2 allele cDNA (SEQ ID NO:11)"
/ FEATURE:
/ NAME/KEY: -
/ LOCATION: 3852..3891
/ OTHER INFORMATION: /note= "start and stop positions for
/ OTHER INFORMATION: genomic sequence surrounding variant
/ OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
/ FEATURE:
/ NAME/KEY: -
/ LOCATION: 5507..6023
/ OTHER INFORMATION: /note= "start and stop positions for
/ OTHER INFORMATION: genomic sequence surrounding variant
/ OTHER INFORMATION: for 24d1(G) allele (SEQ ID NO:20)"
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: replace(3872, "g")
/ OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
/ OTHER INFORMATION: /label= 24d2
/ US-08-834-497A-5

Query Match 100.0%; Score 301; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACCGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCCTCCAAATTC 60
|||||
Db 4570 GGCACCGAATCCCTGGTGGAGTTTCAGAGGTGGCTGAGGCTGTGTGCCCTCCAAATTC 4629

QY 61 TGGGAAGGACTTTCCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 120
|||||
Db 4630 TGGGAAGGACTTTCCTCAATCCCTAGAGTCTCTACCTTATTAATTGAGATGATGAGACAGC 4689

QY 121 CACAAGTCATGGCTTAAATTTCTTCTCCATCATGCTCAAGGGAAGTGTCTATG 180
|||||
Db 4690 CACAAGTCATGGCTTAAATTTCTTCTCCATCATGCTCAAGGGAAGTGTCTATG 4749

QY 181 GCCCTTGCCTTTTATTAAACCAATAATCTTTCTATATTTATACCTGTTAAAAAATTCAGA 240
|||||
Db 4750 GCCCTTGCCTTTTATTAAACCAATAATCTTTCTATATTTATACCTGTTAAAAAATTCAGA 4809

QY 241 AATGTCGAAGCCGGGACGGTGGCTCACCCCTGTAAATCCAGACACTTTGGGAGGCCGAGG 300
|||||
Db 4810 AATGTCGAAGCCGGGACGGTGGCTCACCCCTGTAAATCCAGACACTTTGGGAGGCCGAGG 4869

QY 301 C 301
Db 4870 C 4870

RESULT 8
US-08-834-497A-7
; Sequence 7, Application US/08834497A
; Patent No. 6140305
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gnirke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: HEREDITARY HEMOCHROMATOSIS GENE PRODUCTS
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
```

FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION: /label= 24d1
US-08-834-497A-7

Query Match 100.0%; Score 301; DB 3; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 60
|||||
Db 4570 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGGCTGAGGCTGTGTGCCTCTCCAAATTC 4629
|||||

QY 61 TGGGAGGGAGCTTTCATCTCAGTGTCTACCTTATTAATGAGATGATGAGACAGC 120
|||||
Db 4630 TGGGAGGGAGCTTTCATCTCAGTGTCTACCTTATTAATGAGATGATGAGACAGC 4689
|||||

QY 121 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
|||||
Db 4690 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749
|||||

QY 181 GCCCTTGCTTTTATTTAAACCAATAATCTTTTGTATATTATACCTGTTTAAAAATTCAGA 240
|||||
Db 4750 GCCCTTGCTTTTATTTAAACCAATAATCTTTTGTATATTATACCTGTTTAAAAATTCAGA 4809
|||||

QY 241 AATGTCAAGCGCGGCGACGGTGGCTCACCCCTGTAATCCACGACACTTTGGGAGGCGCGAGG 300
|||||
Db 4810 AATGTCAAGCGCGGCGACGGTGGCTCACCCCTGTAATCCACGACACTTTGGGAGGCGCGAGG 4869
|||||

QY 301 C 301
|

Db 4870 C 4870

RESULT 9
US-09-503-444A-1
Sequence 1, Application US/09503444A
Patent No. 6228594
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503.444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,673

FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..435, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /note= "No. 6228594mal or wild-type (unaffected)"
OTHER INFORMATION: Hereditary Hemochromatosis (HH) gene
OTHER INFORMATION: allele
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: normal or wild-type (unaffected) allele
OTHER INFORMATION: CDNA (SEQ ID NO:9)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: normal or wild-type (unaffected) genomic
OTHER INFORMATION: sequence surrounding variant for 24d2(C)
OTHER INFORMATION: allele (SEQ ID NO:41)"
FEATURE:
NAME/KEY: -
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: normal or wild-type (unaffected) genomic
OTHER INFORMATION: sequence surrounding variant for 24d1(G)
OTHER INFORMATION: allele (SEQ ID NO:20)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "c")
OTHER INFORMATION: /phenotype= "normal or wild-type
OTHER INFORMATION: (unaffected)"
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(3878, "a")
OTHER INFORMATION: /phenotype= "normal or wild-type
OTHER INFORMATION: (unaffected)"
OTHER INFORMATION: /label= 24d7
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "g")
OTHER INFORMATION: /phenotype= "normal or wild-type
OTHER INFORMATION: (unaffected)"
OTHER INFORMATION: /label= 24d1
US-09-503-444A-1

Query Match 100.0%; Score 301; DB 4; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTTCAGAGTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTCAGAGTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629
QY 61 TGGGAAGGACTTTCCTCAATCCTAGAGTCTCTACCTTATAATTTAGATGATGAGACAGC 120
Db 4630 TGGGAAGGACTTTCCTCAATCCTAGAGTCTCTACCTTATAATTTAGATGATGAGACAGC 4689
QY 121 CACAAGTCATGGTTTAAATTTCTTTCTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGTTTAAATTTCTTTCTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749
QY 181 GCCCTTGCCTTTTATTTAAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCCTTGCCTTTTATTTAAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809
QY 241 AATGTCAAGCGCGGCGACGGTGGCTCACCCCTGTAATCCAGCAGCTTTGGGAGCGCGAGG 300
Db 4810 AATGTCAAGCGCGGCGACGGTGGCTCACCCCTGTAATCCAGCAGCTTTGGGAGCGCGAGG 4869
QY 301 C 301
Db 4870 C 4870

RESULT 10

US-09-503-444A-3
; Sequence 3, Application US/09503444A
; Patent No. 6228594
; GENERAL INFORMATION:
; APPLICANT: Thomas, Winston J.
; APPLICANT: Drayna, Dennis T.
; APPLICANT: Feder, John N.
; APPLICANT: Gairke, Andreas
; APPLICANT: Ruddy, David
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Hereditary Hemochromatosis Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect Version 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,444A
; FILING DATE: 14-Feb-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/652,265
; FILING DATE: 23-May-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/632,673
; FILING DATE: 16-Apr-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/630,912
; FILING DATE: 04-Apr-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M.
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 8907-0088-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION: mutation"
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)"
OTHER INFORMATION: gene 24d1 allele"
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: 24d1 allele cDNA (SEQ ID NO:10)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(C) allele (SEQ ID NO:41)"
FEATURE:
NAME/KEY: -
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d1
US-09-503-444A-3
Query Match 100.0%; Score 301; DB 4; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGGAATCCCTGGTTCAGAGTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTCAGAGTTTCAGAGTGGCTGAGGCTGTGTGCTCTCCAAATTC 4629
QY 61 TGGGAAGGACTTTCCTCAATCCTAGAGTCTCTACCTTATAATTTAGATGATGAGACAGC 120
Db 4630 TGGGAAGGACTTTCCTCAATCCTAGAGTCTCTACCTTATAATTTAGATGATGAGACAGC 4689
QY 121 CACAAGTCATGGTTTAAATTTCTTTCTCCATGCATATGGCTCAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGTTTAAATTTCTTTCTCCATGCATATGGCTCAAGGGAAGTGTCTATG 4749
QY 181 GCCCTTGCCTTTTATTTAAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCCTTGCCTTTTATTTAAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 4809
QY 241 AATGTCAAGCGCGGCGAGGCTGGCTCACCCCTGTAATCCAGCAGCTTTGGGAGCGCGAGG 300
Db 4810 AATGTCAAGCGCGGCGAGGCTGGCTCACCCCTGTAATCCAGCAGCTTTGGGAGCGCGAGG 4869
QY 301 C 301
Db 4870 C 4870
RESULT 11
US-09-503-444A-5
; Sequence 5, Application US/09503444A

Patent No. 6228594
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis
OTHER INFORMATION: mutation"
OTHER INFORMATION:
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
OTHER INFORMATION: gene 24d2 allele"
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: 24d2 allele cdna (SEQ ID NO:11)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
FEATURE:
NAME/KEY: -

LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(G) allele (SEQ ID NO:20)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis
OTHER INFORMATION: /label= 24d2
US-09-503-444A-5
Query Match 100.0%; Score 301; DB 4; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGGCTGAGGCTGTCCTCCTCCTCAAAATTC 60
Db 4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGTGGCTGAGGCTGTCCTCCTCCTCAAAATTC 4629
Qy 61 TGGGAAGGGACTTCTCAATCCTAGAGTCTACCTTATATATGAGATGATGAGACAGC 120
Db 4630 TGGGAAGGGACTTCTCAATCCTAGAGTCTACCTTATATATGAGATGATGAGACAGC 4689
Qy 121 CACAAGTCATGGGTTTAAATTTCTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
Db 4690 CACAAGTCATGGGTTTAAATTTCTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749
Qy 181 GCCCTTGGTTTTTAAACCAATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 240
Db 4750 GCCCTTGGTTTTTAAACCAATAATCTTTGTATATTTATACCTGTTAAAAATTCAGA 4809
Qy 241 AATGTCAAGCGCGGCACGGTGGCTCACCCCTGTAATCCACGACACTTTGGGAGGCCGAGG 300
Db 4810 AATGTCAAGCGCGGCACGGTGGCTCACCCCTGTAATCCACGACACTTTGGGAGGCCGAGG 4869
Qy 301 C 301
Db 4870 C 4870
RESULT 12
US-09-503-444A-7
Sequence 7, Application US/09503444A
Patent No. 6228594
GENERAL INFORMATION:
APPLICANT: Thomas, Winston J.
APPLICANT: Drayna, Dennis T.
APPLICANT: Feder, John N.
APPLICANT: Gnirke, Andreas
APPLICANT: Ruddy, David
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Hereditary Hemochromatosis Gene
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect Version 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/503,444A
FILING DATE: 14-Feb-2000
CLASSIFICATION:
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/652,265
FILING DATE: 23-May-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/632,673
FILING DATE: 16-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,912
FILING DATE: 04-Apr-1996
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0088-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 10825 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: Join(361..436, 3762..4025, 4235..4510, 5606..5881,
LOCATION: 6040..6153, 7107..7147)
OTHER INFORMATION: /product= "Hereditary Hemochromatosis"
OTHER INFORMATION: and 24d2 mutations"
OTHER INFORMATION: /note= "Hereditary Hemochromatosis (HH)
OTHER INFORMATION: gene containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles"
FEATURE:
NAME/KEY: -
LOCATION: 140..7319
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: cDNA containing a combination of both
OTHER INFORMATION: 24d1 and 24d2 alleles
OTHER INFORMATION: (SEQ ID NO:12)"
FEATURE:
NAME/KEY: -
LOCATION: 3852..3891
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d2(G) allele (SEQ ID NO:42)"
FEATURE:
NAME/KEY: -
LOCATION: 5507..6023
OTHER INFORMATION: /note= "start and stop positions for
OTHER INFORMATION: genomic sequence surrounding variant
OTHER INFORMATION: for 24d1(A) allele (SEQ ID NO:21)"
FEATURE:
NAME/KEY: allele
LOCATION: replace(3872, "g")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d2
FEATURE:
NAME/KEY: allele
LOCATION: replace(5834, "a")
OTHER INFORMATION: /phenotype= "Hereditary Hemochromatosis"
OTHER INFORMATION:
OTHER INFORMATION: /label= 24d1
US-09-503-444A-7
Query Match 100.0%; Score 301; DB 4; Length 10825;
Best Local Similarity 100.0%; Pred. No. 6.6e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCCTCTCCAAATTC 60
|

Db 4570 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCCTCTCCAAATTC 4629
Qy 61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
|
Db 4630 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 4689
Qy 121 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
|
Db 4690 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 4749
Qy 181 GCCCTTGCTTTTATTAAACCAATAAATCTTTTGTATATTATACCTCTTAAAAATTCAGA 240
|
Db 4750 GCCCTTGCTTTTATTAAACCAATAAATCTTTTGTATATTATACCTCTTAAAAATTCAGA 4809
Qy 241 AATGTCAAGCGCGGCACGGTGGCTCACCCCTGTAATCCAGCAGCTTTGGGAGGCCGAGG 300
|
Db 4810 AATGTCAAGCGCGGCACGGTGGCTCACCCCTGTAATCCAGCAGCTTTGGGAGGCCGAGG 4869
Qy 301 C 301
|
Db 4870 C 4870
RESULT 13
US-09-277-457-27
; Sequence 27, Application US/09277457
; Patent No. 6355425
; GENERAL INFORMATION:
; APPLICANT: Rothenberg, Barry E.
; APPLICANT: Sawada-Hirai, Ritsuko
; APPLICANT: Barton, James C.
; TITLE OF INVENTION: MUTATIONS ASSOCIATED WITH IRON DISORDERS
; FILE REFERENCE: 10653/002001
; CURRENT APPLICATION NUMBER: US/09/277,457
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 12146
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-277-457-27
Query Match 100.0%; Score 301; DB 4; Length 12146;
Best Local Similarity 100.0%; Pred. No. 6.9e-86;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGCACGGAATCCCTGGTTGGAGTTTCAGAGGTGGCTGAGGCTGTGTCCTCTCCAAATTC 60
|
Db 5460 ggcacggaatccctgggttggagtttcagaggtggctgaggtgctgctctccaaattc 5519
Qy 61 TGGGAAGGGACTTTCTCAATCCTAGAGTCTCTACCTTATAATTGAGATGTATGAGACAGC 120
|
Db 5520 tgggaaggagacttttcaatcctcagtagtctctacattataattgagatgtatgagacagc 5579
Qy 121 CACAAGTCATGGGTTTAATTTCTTTCTCCATGCATATGGCTCAAAGGGAAGTGTCTATG 180
|
Db 5580 cacaagtcatgggtttaattttcttctccatgcataatggctcaagggaagtgtctatg 5639
Qy 181 GCCCTTGCTTTTATTAAACCAATAAATCTTTTGTATATTATACCTCTTAAAAATTCAGA 240
|
Db 5640 gcccttgctttttatttaaccaataatttttggatatattatactgttaaaaattcaga 5699
Qy 241 AATGTCAAGCGCGGCACGGTGGCTCACCCCTGTAATCCAGCAGCTTTGGGAGGCCGAGG 300
|
Db 5700 aatgtcaaggcgggcacggtggctcaccctgttaatcccgagcactttgggagggccgagg 5759
Qy 301 C 301
|
Db 5760 c 5760

```
RESULT 14
US-08-724-394A-20
; Sequence 20, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match 100.0%; Score 301; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 2.5e-85;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGCGCTGAGGCTGCTCCAAATTC 60
|||||
Db 196873 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGCGCTGAGGCTGCTCCAAATTC 196932
|||||

QY 61 TGGGAAGGACTTCTCAATCCTAGAGTCTACCTTATAATTGAGATGTATGAGACAGC 120
|||||
Db 196993 CACAAGTCATGGTTTAAATTTCTTCCATGATATGGCTCAAGGAAAGTGTCTATG 197052
|||||

QY 181 GCCCTTGCTTTTATTTAAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 240
|||||
Db 197053 GCCCTTGCTTTTATTTAAACCAATAATCTTTTGTATATTTATACCTGTTAAAAATTCAGA 197112
|||||

QY 241 RATGTCAAGGGCGGCGAGGTTGGCTCACCCTGTAAATCCAGACACTTTGGGAGGCCGAGG 300
|||||
```

```
Db 197113 AATGTCAAGGGCGGCGAGGTTGGCTCACCCTGTAAATCCAGACACTTTGGGAGGCCGAGG 197172
|||||

QY 301 C 301
|
Db 197173 C 197173

RESULT 15
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match 100.0%; Score 301; DB 2; Length 246240;
Best Local Similarity 100.0%; Pred. No. 2.5e-85;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGCGCTGAGGCTGCTCCAAATTC 60
|||||
Db 196873 GGCACGGAATCCCTGGTGGAGTTTCAGAGTGCGCTGAGGCTGCTCCAAATTC 196932
|||||

QY 61 TGGGAAGGACTTCTCAATCCTAGAGTCTACCTTATAATTGAGATGTATGAGACAGC 120
|||||
Db 196933 TGGGAAGGACTTCTCAATCCTAGAGTCTACCTTATAATTGAGATGTATGAGACAGC 196992
|||||

QY 121 CACAAGTCATGGGTTTAAATTTCTTCCATGATATGGCTCAAGGAAAGTGTCTATG 180
|||||
```

```
|||||
Db 196993 CACAAGTCATGGGTTTAAATTTCTTTCTCCATGCAATAGGCTAAAGGGAAGTCTCTATG 197052
Qy 181 GCCCTTGCTTTTATTAAACCAATAATCTTTTGTATATTATACCTGTTAAAAAATTCAGA 240
|||||
Db 197053 GCCCTTGCTTTTATTAAACCAATAATCTTTTGTATATTATACCTGTTAAAAAATTCAGA 197112
Qy 241 AATGTCAAGCGCGGCACGGTGGCTCACCCCTGTAAATCCAGCACACTTTGGGAGGCCGAGG 300
|||||
Db 197113 AATGTCAAGCGCGGCACGGTGGCTCACCCCTGTAAATCCAGCACACTTTGGGAGGCCGAGG 197172
Qy 301 C 301
Db 197173 C 197173
```

Search completed: June 19, 2002, 10:58:07
Job time: 8225 sec

AAACTTTGCCACATGTCACCTAGTAGACAAACTCCTCGTTAAGAAAGCTCGGGTTCAAAAAATTAACAAGTA 2170 2180 2190 2200 2210 2220 2230
2240 2250 2260 2270 2280 2290 2300
GTGCTGGGAGTAGAGGCCAAGAAGTATAGTGGCTCAGAGAGGAGCCACAAACAGGTTGTCAGGGG
|||||
GTGCTGGGAGTAGAGGCCAAGAAGTATAGTGGCTCAGAGAGGAGCCACAAACAGGTTGTCAGGGG 2240 2250 2260 2270 2280 2290 2300
2310 2320 2330 2340 2350 2360 2370
CCTGTAGGCTGTGGTGAATTTCTAGCCAAAGGAGTAACAGTATCTGTACAGGCTTTTAAAGATTCTCT
CCTGTAGGCTGTGGTGAATTTCTAGCCAAAGGAGTAACAGTATCTGTACAGGCTTTTAAAGATTCTCT 2310 2320 2330 2340 2350 2360 2370
2380 2390 2400 2410 2420 2430 2440
GCTGCTATGTGGAAGCAGAAATGAAGGAGCAACAGTAAGAGAGGAGCCACAGCAAGGCTGTATAC
|||||
GCTGCTATGTGGAAGCAGAAATGAAGGAGCAACAGTAAGAGAGGAGCCACAGCAAGGCTGTATAC 2380 2390 2400 2410 2420 2430 2440
2450 2460 2470 2480 2490 2500 2510 2520
ACTTCCAGGCAAGAGTAGTGGAGTGGCTGGTGGGAACAGAAAGGAGGTGACAAACATTGTCTCCTGAA
|||||
AGTCCAGGCAAGAGTAGTGGAGTGGCTGGTGGGAACAGAAAGGAGGTGACAAACATTGTCTCCTGAA 2450 2460 2470 2480 2490 2500 2510 2520
TATATTCTCAAGGAAGTTCTCTCAAGGATTCTATGTTGTGTAGAGAAAGAGAAGATTGGCTGGGTGAGTA 2530 2540 2550 2560 2570 2580 2590
TATATTCTCAAGGAAGTTCTCTCAAGGATTCTATGTTGTGTAGAGAAAGAGAAGATTGGCTGGGTGAGTA 2530 2540 2550 2560 2570 2580 2590
2600 2610 2620 2630 2640 2650 2660
GCTCATGCCAAGGAGGCCAAGAGAGACAGATTCTCGAGCTCAGGAGTTCAAGACACAGCTGGGCAACAC
GCTCATGCCAAGGAGGCCAAGAGAGACAGATTCTCGAGCTCAGGAGTTCAAGACACAGCTGGGCAACAC 2600 2610 2620 2630 2640 2650 2660
2670 2680 2690 2700 2710 2720 2730
AGCAAAACCCCTTCTCAAAAAATACAAAAATTAGCTGGGTGTGGTGGACCTGTGATCCTAGCTA
AGCAAAACCCCTTCTCAAAAAATACAAAAATTAGCTGGGTGTGGTGGACCTGTGATCCTAGCTA 2670 2680 2690 2700 2710 2720 2730
2740 2750 2760 2770 2780 2790 2800
CTCGGAGGCTGAGGTGAGGTATTGCTTGAGCCAGGAAGTTGAGGCTGAGTGAGCCATGACTGTGCCA
CTCGGAGGCTGAGGTGAGGTATTGCTTGAGCCAGGAAGTTGAGGCTGAGTGAGCCATGACTGTGCCA 2740 2750 2760 2770 2780 2790 2800
2810 2820 2830 2840 2850 2860 2870 2880
CTGTACTTCAGCCTAGGTGACAGAGCAAGCCCTGTCTCCCTGACCCCTGAAAAAGAGAGATTAAAGT
CTGTACTTCAGCCTAGGTGACAGAGCAAGCCCTGTCTCCCTGACCCCTGAAAAAGAGAGATTAAAGT 2810 2820 2830 2840 2850 2860 2870 2880
2890 2900 2910 2920 2930 2940 2950
TGACTTTGTTCTTTATTTATTTGGCTGAGCAGTGGGTAATTGGCAATGCCATTCTCAGATGGT
TGACTTTGTTCTTTATTTATTTGGCTGAGCAGTGGGTAATTGGCAATGCCATTCTCAGATGGT 2890 2900 2910 2920 2930 2940 2950
2960 2970 2980 2990 3000 3010 3020
GAAGGCAGAGGAAGAGAGTTGGGGTAAATCAAGGATCTGCATTGGGACATTGAAGTTGAGATTCCA
GAAGGCAGAGGAAGAGAGTTGGGGTAAATCAAGGATCTGCATTGGGACATTGAAGTTGAGATTCCA 2960 2970 2980 2990 3000 3010 3020
3030 3040 3050 3060 3070 3080 3090
GTGAGCTTCCAAAGTGGTGGGCCACATAGGAGTTGAGTTAAGAAATTCAGGACCAAGGCTGGCAGGGTG 3030 3040 3050 3060 3070 3080 3090
GTCAGGCTTCCAAAGTGGTGGGCCACATAGGAGTTGAGTTAAGAAATTCAGGACCAAGGCTGGCAGGGTG 3030 3040 3050 3060 3070 3080 3090
GTCAGGCTTCCAAAGTGGTGGGCCACATAGGAGTTGAGTTAAGAAATTCAGGACCAAGGCTGGCAGGGTG 3030 3040 3050 3060 3070 3080 3090

3030 3040 3050 3060 3070 3080 3090
3100 3110 3120 3130 3140 3150 3160
GCTCACATTCTGTAATCCACACTTTGGTGGCTGAGCAGGTAGATCATTTTGGTTCAGAGTTTCGAGACAA
|||||
GCTCACATTCTGTAATCCACACTTTGGTGGCTGAGCAGGTAGATCATTTTGGTTCAGAGTTTCGAGACAA 3100 3110 3120 3130 3140 3150 3160
3170 3180 3190 3200 3210 3220 3230 3240
GCTTGGCCAAATGGTGAACCCCATGTCTTACTTAAATAACAAAAATTAGCCTGGTGTGGCGCACGCCT
GCTTGGCCAAATGGTGAACCCCATGTCTTACTTAAATAACAAAAATTAGCCTGGTGTGGCGCACGCCT 3170 3180 3190 3200 3210 3220 3230 3240
3250 3260 3270 3280 3290 3300 3310
ATAGTCCACAGGTTTTCAGGAGGCTTAGGTAGGAGATCCCTTGAACCCAGGAGGTGCAGGTTTGCAGTGAGCT
ATAGTCCACAGGTTTTCAGGAGGCTTAGGTAGGAGATCCCTTGAACCCAGGAGGTGCAGGTTTGCAGTGAGCT 3250 3260 3270 3280 3290 3300 3310
3320 3330 3340 3350 3360 3370 3380
GAGATTCTGCCACTGCACTCCAGCTCCAGCTGGGTGATAGGTGAGACTCTGTCTCAAAAAATAAAAAA
|||||
GAGATTCTGCCACTGCACTCCAGCTCCAGCTGGGTGATAGGTGAGACTCTGTCTCAAAAAATAAAAAA 3320 3330 3340 3350 3360 3370 3380
3390 3400 3410 3420 3430 3440 3450
AAAAAAACTGAAGGAATTTCTCTCAGGATTTGGGTCTAATTTGGCCTGAGCACCACCTCTCTGAGTTCAAC
AAAAAAACTGAAGGAATTTCTCTCAGGATTTGGGTCTAATTTGGCCTGAGCACCACCTCTCTGAGTTCAAC 3390 3400 3410 3420 3430 3440 3450
3460 3470 3480 3490 3500 3510 3520
TACCATGGCTAGACACACCTTAACATTTCTAGAATCCACAGCTTTAGTGAGTCTGTCTAATCATGAGTA
TACCATGGCTAGACACACCTTAACATTTCTAGAATCCACAGCTTTAGTGAGTCTGTCTAATCATGAGTA 3460 3470 3480 3490 3500 3510 3520
3530 3540 3550 3560 3570 3580 3590 3600
TTGGAATAGGATCTGGGGGAGTGGGGGTGGCAGCCAGTGTGGCAGAGAAAGCACAAAGGAAAGAGC
TTGGAATAGGATCTGGGGGAGTGGGGGTGGCAGCCAGTGTGGCAGAGAAAGCACAAAGGAAAGAGC 3530 3540 3550 3560 3570 3580 3590 3600
3610 3620 3630 3640 3650 3660 3670
ACCCAGAGCTGTCTATATGGAAGAAAGACAGGACTGCAACTCACCCCTTCAAAAAATGAGGACCAGACAGCT
ACCCAGAGCTGTCTATATGGAAGAAAGACAGGACTGCAACTCACCCCTTCAAAAAATGAGGACCAGACAGCT 3610 3620 3630 3640 3650 3660 3670
3680 3690 3700 3710 3720 3730 3740
GATGATGAGTTGATCAGGTGTGGAGCCTCAACATCCTGCTCCCTCTACTACATGTTAAGGCC
GATGATGAGTTGATCAGGTGTGGAGCCTCAACATCCTGCTCCCTCTACTACATGTTAAGGCC 3680 3690 3700 3710 3720 3730 3740
3750 3760 3770 3780 3790 3800 3810
TGTTGCTGTCTCCAGTTCCAGTTCCACTCTCTGCACTACTCTTCTATGGGTGCTCAGAGAGGACCTTGGTCT
TGTTGCTGTCTCCAGTTCCAGTTCCACTCTCTGCACTACTCTTCTATGGGTGCTCAGAGAGGACCTTGGTCT 3750 3760 3770 3780 3790 3800 3810
3820 3830 3840 3850 3860 3870 3880
TTCTTGTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTCTCTGTGTTCTATGATNATGAGTCCGCTGT
TTCTTGTGTTTGAAGCTTTGGGCTACGTGGATGACCAGCTGTCTCTGTGTTCTATGATNATGAGTCCGCTGT 3820 3830 3840 3850 3860 3870 3880
3890 3900 3910 3920 3930 3940 3950 3960
GGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAAGTGCAGAGTCTGAA
GGAGCCCGAACTCCATGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGCTGCAAGTGCAGAGTCTGAA 3890 3900 3910 3920 3930 3940 3950 3960

AAATCACCATGAAGTGGCTGAAGGATGAAGCAGCCAAATGGATGCCAAGGAGTTGGAACCTAAAGACGATATG
|||||
AAATCACCATGAAGTGGCTGAAGGATGAAGCAGCCAAATGGATGCCAAGGAGTTGGAACCTAAAGACGATATG
5690 5700 5710 5720 5730 5740 5750 5760
CCCAATGGGATGGACCTACAGGGCTGGGATGAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATAG
|||||
CCCAATGGGATGGACCTACAGGGCTGGGATGAACCTTGGCTGTACCCCTGGGGAAGAGCAGAGATATAG
5770 5780 5790 5800 5810 5820 5830
5840 5850 5860 5870 5880 5890 5900
TNCAGGTGGAGCACCAGGCTGGATGAGCCCTCATTTGGTGTACCTGGGATGTGACTGATGAGAGCCAGG
|||||
TACAGGTGGAGCACCAGGCTGGATGAGCCCTCATTTGGTGTACCTGGGATGTGACTGATGAGAGCCAGG
5840 5850 5860 5870 5880 5890 5900
5910 5920 5930 5940 5950 5960 5970
AGCTGAGAAATCTATTTGGGGTTGAGAGGAGTGCCTGAGAGGTAATTAATGGCAGTGAGATGAGGATGCG
|||||
AGCTGAGAAATCTATTTGGGGTTGAGAGGAGTGCCTGAGAGGTAATTAATGGCAGTGAGATGAGGATGCG
5910 5920 5930 5940 5950 5960 5970
5980 5990 6000 6010 6020 6030 6040
TCTTTGTAGGGTGGCTGAGGTCGCAATCAAGGCTTTAACTTGTCTTTTGTATGAGCCCTCAC
|||||
TCTTTGTAGGGTGGCTGAGGTCGCAATCAAGGCTTTAACTTGTCTTTTGTATGAGCCCTCAC
5980 5990 6000 6010 6020 6030 6040
6050 6060 6070 6080 6090 6100 6110 6120
CGCTGGCACCCTAGTCAATCGAGTCATCACTGGAATGCTGTTTGTCTGCAATCTGTTCAATGGAATTT
|||||
CGCTGGCACCCTAGTCAATCGAGTCATCACTGGAATGCTGTTTGTCTGCAATCTGTTCAATGGAATTT
6050 6060 6070 6080 6090 6100 6110 6120
6130 6140 6150 6160 6170 6180 6190
TGTTCAATATTAAGAAGAGCAGGTTCAAGTGAAGTGAAGACAGGGGAGTCTCTAGTACCTGCG
|||||
TGTTCAATATTAAGAAGAGCAGGTTCAAGTGAAGTGAAGACAGGGGAGTCTCTAGTACCTGCG
6130 6140 6150 6160 6170 6180 6190
6200 6210 6220 6230 6240 6250 6260
CCAGGACACAGTGGAGAGGAGGAGTCTGGCATCCATCGAAGTGAAGTCTCTCAATTTATTTCT
|||||
CCAGGACACAGTGGAGAGGAGGAGTCTGGCATCCATCGAAGTGAAGTCTCTCAATTTATTTCT
6200 6210 6220 6230 6240 6250 6260
6270 6280 6290 6300 6310 6320 6330
TTGGGACACAGCAGTCCCTGGGAGACAGAAATTAATGGTCTCCCGAATGAAGTCTCTAATTCAC
|||||
TTGGGACACAGCAGTCCCTGGGAGACAGAAATTAATGGTCTCCCGAATGAAGTCTCTAATTCAC
6270 6280 6290 6300 6310 6320 6330
6340 6350 6360 6370 6380 6390 6400
AAACATCTCAGACCACTACTATTTTGAAGAGCTGTTTAAAGTAGTACAGGGCTTTGAGTCTGAGAAT
|||||
AAACATCTCAGACCACTACTATTTTGAAGAGCTGTTTAAAGTAGTACAGGGCTTTGAGTCTGAGAAT
6340 6350 6360 6370 6380 6390 6400
6410 6420 6430 6440 6450 6460 6470 6480
CACTGGCTATCTCAGAACCCAAATCTGGTAGGGAATGAATTTGATAGCAAGTAATGTAGTTAAAGAAG
|||||
CACTGGCTATCTCAGAACCCAAATCTGGTAGGGAATGAATTTGATAGCAAGTAATGTAGTTAAAGAAG
6410 6420 6430 6440 6450 6460 6470 6480
6490 6500 6510 6520 6530 6540 6550
ACCCATGAGGTCCTAAGCAGGAGCAATGCTTAGGGTGTCAAGGAAGAATGATCACAATTCAGC
|||||
ACCCATGAGGTCCTAAGCAGGAGCAATGCTTAGGGTGTCAAGGAAGAATGATCACAATTCAGC
6490 6500 6510 6520 6530 6540 6550
6560 6570 6580 6590 6600 6610 6620
TGGGATCAAGATAGCCTTCTGGATCTTGAAGGAGAGCTGGATTCCATTAGGTGAGGTGAAGATGATGG

|||||
TGGGATCAAGATAGCCTTCTGATCTTTGAAGGAGAGCTGGATTCCATTAGGTGAGGTGAAGATGATGGG
6560 6570 6580 6590 6600 6610 6620
6630 6640 6650 6660 6670 6680 6690
AGGTCTACAGACGAGCAACCATGCAAGTAGGAGATATAAGCATATTCACCCCTGAACATCTGGT
|||||
AGGTCTACAGACGAGCAACCATGCAAGTAGGAGATATAAGCATATTCACCCCTGAACATCTGGT
6630 6640 6650 6660 6670 6680 6690
6700 6710 6720 6730 6740 6750 6760
TGACTCTTAACCTGAGTTGCTAGCTATCACTACCAATTTGCAATTTACCCCTGAACATCTGGT
|||||
TGACTCTTAACCTGAGTTGCTAGCTATCACTACCAATTTGCAATTTACCCCTGAACATCTGGT
6700 6710 6720 6730 6740 6750 6760
6770 6780 6790 6800 6810 6820 6830 6840
GTAGGAAAGAGCAATCAGAAAGCCAGCTCATACAGATCCAAGGTCCTTTGGGATATTTGGTTATGA
|||||
GTAGGAAAGAGCAATCAGAAAGCCAGCTCATACAGATCCAAGGTCCTTTGGGATATTTGGTTATGA
6770 6780 6790 6800 6810 6820 6830 6840
6850 6860 6870 6880 6890 6900 6910
TCACTGGGTGTCATTGAAGGATCCCTAAGAAAGGAGGACCATCTCCCTTATATGTTGTTTGA
|||||
TCACTGGGTGTCATTGAAGGATCCCTAAGAAAGGAGGACCATCTCCCTTATATGTTGTTTGA
6850 6860 6870 6880 6890 6900 6910
6920 6930 6940 6950 6960 6970 6980
AGAAGTTAGATGAGAGTGGAGACCAAGTGAAGAAAGCAATTAAGCATTTCCAGATGAGAGATATGTTCT
|||||
AGAAGTTAGATGAGAGTGGAGACCAAGTGAAGAAAGCAATTAAGCATTTCCAGATGAGAGATATGTTCT
6920 6930 6940 6950 6960 6970 6980
6990 7000 7010 7020 7030 7040 7050
TGAATCCCAATAGTGGCCAGGTCATAAATTGAGATGGTGAANTGAGGAAATTAAGAAAGAGAGGCAAG
|||||
TGAATCCCAATAGTGGCCAGGTCATAAATTGAGATGGTGAANTGAGGAAATTAAGAAAGAGAGGCAAG
6990 7000 7010 7020 7030 7040 7050
7060 7070 7080 7090 7100 7110 7120
ATGGTCCCTAGTTTGTGATGCTCTTCTGGTCTCTGTCTCCACAGGAGGACCATGGGACACACGT
|||||
ATGGTCCCTAGTTTGTGATGCTCTTCTGGTCTCTGTCTCCACAGGAGGACCATGGGACACACGT
7060 7070 7080 7090 7100 7110 7120
7130 7140 7150 7160 7170 7180 7190 7200
CTTAGTGAACGTGAGTGACACGACCTGCGAGACTCACTGTGGGAGGAGACAACTAGAGACTCAAGA
|||||
CTTAGTGAACGTGAGTGACACGACCTGCGAGACTCACTGTGGGAGGAGACAACTAGAGACTCAAGA
7130 7140 7150 7160 7170 7180 7190 7200
7210 7220 7230 7240 7250 7260 7270
GGAGTGCATTTAGCTTCTGTTCAATTTCCCAAGAGTGAACCTAAACATAGAAATTCCTGACCAACT
|||||
GGAGTGCATTTAGCTTCTGTTCAATTTCCCAAGAGTGAACCTAAACATAGAAATTCCTGACCAACT
7210 7220 7230 7240 7250 7260 7270
7280 7290 7300 7310 7320 7330 7340
CCTTGATTTAGCTTCTGTTCAATTTCCCAAGAGTGAACCTAAACATAGAAATTCCTGACCAACT
|||||
CCTTGATTTAGCTTCTGTTCAATTTCCCAAGAGTGAACCTAAACATAGAAATTCCTGACCAACT
7280 7290 7300 7310 7320 7330 7340
7350 7360 7370 7380 7390 7400 7410
CGGTGATCCCTAGCTTCTGACCTTCCCTGGAACCTCTCTCATGAACCTCAAGCTGATCTAGAGCTGCC
|||||
CGGTGATCCCTAGCTTCTGACCTTCCCTGGAACCTCTCTCATGAACCTCAAGCTGATCTAGAGCTGCC
7350 7360 7370 7380 7390 7400 7410
7420 7430 7440 7450 7460 7470 7480
TTTCATTTCTCGTCCCTCAGACATACACCTATGTCATTTTCATTTTCATTTTGAAGAGGACTCCTT
|||||

TTTCAATTCCTCGGTACCTTCACGACATACACCTATGTCTATTTCATTTTCCTTATTTTGGAAAGAGACTCCTT
7420 7430 7440 7450 7460 7470 7480
7490 7500 7510 7520 7530 7540 7550 7560
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAGCTTTGAACCCCTGGGACCTGGCTAGTCAAT
TT
AAATTTGGGGGACTTACATGATTCATTTTAACATCTGAGAAAGCTTTGAACCCCTGGGACCTGGCTAGTCAAT
7490 7500 7510 7520 7530 7540 7550 7560
7570 7580 7590 7600 7610 7620 7630
AACCTTACCAGATTTTTTACACATGATCTATGCATTTTCTGGACCCGTTCAACTTTTCCCTTGAATCCTTC
TT
AACCTTACCAGATTTTTTACACATGATCTATGCATTTTCTGGACCCGTTCAACTTTTCCCTTGAATCCTTC
7570 7580 7590 7600 7610 7620 7630
7640 7650 7660 7670 7680 7690 7700
TCTGTGTTACCCAGTAACATCTCTGTACCAGCCTTGGGGATTTCTCCATCTGATTTGATGTGAGTTGCA
TCTGTGTTACCCAGTAACATCTCTGTACCAGCCTTGGGGATTTCTCCATCTGATTTGATGTGAGTTGCA
7640 7650 7660 7670 7680 7690 7700
7710 7720 7730 7740 7750 7760 7770
CAGCTATGAAGGCTGTACACTGCACGAATGGAAGGACCTTGTCCAGAAAAAGCATCATGGCTATCTGTG
TT
7710 7720 7730 7740 7750 7760 7770
7780 7790 7800 7810 7820 7830 7840
GGTAGTATCATGGGTGTTTTAGCAGGTAGGAGCAATATCTTGAAGGGTGTCTGAAGGGTGTGTTTTTC
TT
GGTAGTATCATGGGTGTTTTAGCAGGTAGGAGCAATATCTTGAAGGGTGTCTGAAGGGTGTGTTTTTC
7780 7790 7800 7810 7820 7830 7840
7850 7860 7870 7880 7890 7900 7910 7920
TAATTTGGCATGAAGGTGTATACAGATTTTGAAGTTTAAATGGTGCCTTCAATTTGGGATGCTACTCTAGTAT
TT
TAATTTGGCATGAAGGTGTATACAGATTTTGAAGTTTAAATGGTGCCTTCAATTTGGGATGCTACTCTAGTAT
7850 7860 7870 7880 7890 7900 7910 7920
7930 7940 7950 7960 7970 7980 7990
TCCAGACCTGAGAATCACAAATTTTCTACTCGTCTCTCTTCTGATATGAAATTAATGATAAGG
TT
TCCAGACCTGAGAATCACAAATTTTCTACTCGTCTCTCTTCTGATATGAAATTAATGATAAGG
7930 7940 7950 7960 7970 7980 7990
8000 8010 8020 8030 8040 8050 8060
ATGATAAAGCACTTACTTCGTGTCGAGCTCTTCTGAGCACCTACTTACATGCATTACTGCAATGCACTTCTT
TT
ATGATAAAGCACTTACTTCGTGTCGAGCTCTTCTGAGCACCTACTTACATGCATTACTGCAATGCACTTCTT
8000 8010 8020 8030 8040 8050 8060
8070 8080 8090 8100 8110 8120 8130
ACAATAATTCATGAGATAGGTACTATATCCCATTTCTTTTAAATGAAGAAAGTGAAGTAGGCCGGGC
TT
ACAATAATTCATGAGATAGGTACTATATCCCATTTCTTTTAAATGAAGAAAGTGAAGTAGGCCGGGC
8070 8080 8090 8100 8110 8120 8130
8140 8150 8160 8170 8180 8190 8200
ACGGTGGCTCACGCCGTGAATCCAGCACTTTGGGAGGCCAAGCGGGTGGATCACAGGCTCAGGAGATCGA
TT
ACGGTGGCTCACGCCGTGAATCCAGCACTTTGGGAGGCCAAGCGGGTGGATCACAGGCTCAGGAGATCGA
8140 8150 8160 8170 8180 8190 8200
8210 8220 8230 8240 8250 8260 8270 8280
GACCATCTCGCTTAACATGGTGAACCCCATCTCTAATAAATAACAAAAATAGCTGGCGTGGTGGCAG
TT
GACCATCTCGCTTAACATGGTGAACCCCATCTCTAATAAATAACAAAAATAGCTGGCGTGGTGGCAG
8210 8220 8230 8240 8250 8260 8270 8280
8290 8300 8310 8320 8330 8340 8350
A2GGCTGTAGTCCAGCTACTCGGAAGGCTGAGGCAGGAGATGGCATGAACCCAGGAGGCAGAGCTTCGAG
TT
ACGCTGTAGTCCAGCTACTCGGAAGGCTGAGGCAGGAGATGGCATGAACCCAGGAGGCAGAGCTTCGAG

8360 8370 8380 8390 8400 8410 8420
TGAGCCAGTTTGGCCACTGCACTCCAGCCTAGGTGACAGAGTGAGACTPCCATCTCAAAAAAATAAATA
TT
TGAGCCAGTTTGGCCACTGCACTCCAGCCTAGGTGACAGAGTGAGACTPCCATCTCAAAAAAATAAATA
8360 8370 8380 8390 8400 8410 8420
8430 8440 8450 8460 8470 8480 8490
AAAAATAAAAAATGAAAAAAGAAAGTGAAGTATAGAGTATCTCATAGTTTGTCACTGATAGAAACAGG
TT
AAAAATAAAAAATGAAAAAAGAAAGTGAAGTATAGAGTATCTCATAGTTTGTCACTGATAGAAACAGG
8430 8440 8450 8460 8470 8480 8490
8500 8510 8520 8530 8540 8550 8560
TTTCAAACTCAGTCAATCTCACCGTTTGATACATCTCAGACACCCTACATTCAGTAGTTTAGATCCCTAGA
TT
TTTCAAACTCAGTCAATCTCACCGTTTGATACATCTCAGACACCCTACATTCAGTAGTTTAGATCCCTAGA
8500 8510 8520 8530 8540 8550 8560
8570 8580 8590 8600 8610 8620 8630 8640
ATAAATAGAGAAGGAAGAGATGGCTCTTCTGTCTCATTTGTGTTCTTCTGAGTGAGCTTGAATACAT
TT
ATAAATAGAGAAGGAAGAGATGGCTCTTCTGTCTCATTTGTGTTCTTCTGAGTGAGCTTGAATACAT
8570 8580 8590 8600 8610 8620 8630 8640
8650 8660 8670 8680 8690 8700 8710
GAAGGGAAACAGCAGAAAAACAACACTGATCTCAGCTGTCTCATGTTTCTTAAAGTCCCTGAAGGAAGG
TT
GAAGGGAAACAGCAGAAAAACAACACTGATCTCAGCTGTCTCATGTTTCTTAAAGTCCCTGAAGGAAGG
8650 8660 8670 8680 8690 8700 8710
8720 8730 8740 8750 8760 8770 8780
TCCGTGAATGTGACTCCTCTGCTCTGTGCTCTCTTTGGCAATTCATTTCTTGGACCTACGCAAGGAC
TT
TCCGTGAATGTGACTCCTCTGCTCTGTGCTCTCTTTGGCAATTCATTTCTTGGACCTACGCAAGGAC
8720 8730 8740 8750 8760 8770 8780
8790 8800 8810 8820 8830 8840 8850
TGTAATTTGGTGGGACAGCTAGTGGCCCTGTGGCTTACACACGGTGTCTCCCTAGGCCCTGCTCTG
TT
TGTAATTTGGTGGGACAGCTAGTGGCCCTGTGGCTTACACACGGTGTCTCCCTAGGCCCTGCTCTG
8790 8800 8810 8820 8830 8840 8850
8860 8870 8880 8890 8900 8910 8920
GAGTCAGAACTCTCGTGGTATTTCCCTCAATGAAGTGGAGTAAAGCTCTCATTTTGGATGGTAAATGGA
TT
GAGTCAGAACTCTCGTGGTATTTCCCTCAATGAAGTGGAGTAAAGCTCTCATTTTGGATGGTAAATGGA
8860 8870 8880 8890 8900 8910 8920
8930 8940 8950 8960 8970 8980 8990 9000
AGCCACCAAGTGGCTTAGAGAGATGCCAGGTCCCTCCATGGAGCCACTGGGGTCCGGTGCAACATTAATAA
TT
AGCCACCAAGTGGCTTAGAGAGATGCCAGGTCCCTCCATGGAGCCACTGGGGTCCGGTGCAACATTAATAA
8930 8940 8950 8960 8970 8980 8990 9000
9010 9020 9030 9040 9050 9060 9070
AAAACTAACCCAGACATTCAGGAATTTGCTAGATTTCTGGAAATCAGTTTCCACCATGTTTCAAAAGAGTCTTTT
9010 9020 9030 9040 9050 9060 9070
9080 9090 9100 9110 9120 9130 9140
TTTTTTTTTTGAGACTCTATTTGCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGTAACTCTGCCCT
TT
TTTTTTTTTTGAGACTCTATTTGCCAGGCTGGAGTGCAATGGCATGATCTCGGCTCACTGTAACTCTGCCCT
9080 9090 9100 9110 9120 9130 9140
9150 9160 9170 9180 9190 9200 9210
CCGAGGTTCAAGCGATTTCTCTGTCTACGCTCCCAAGTAGCTGGGATTACAGCGGTGCACCGATGCCCG
TT
CCGAGGTTCAAGCGATTTCTCTGTCTACGCTCCCAAGTAGCTGGGATTACAGCGGTGCACCGATGCCCG
9150 9160 9170 9180 9190 9200 9210

X 10 20 30 40 50 60 70
TCTAAGGTGAGATAAAATTTTAAATGATGATTTGAATTTTGAATTCATAAATATTTAAATATCTAAAGT
TCTAAGGTGAGATAAAATTTTAAATGATGATTTGAATTTTGAATTCATAAATATTTAAATATCTAAAGT
X 10 20 30 40 50 60 70
80 90 100 110 120 130 140
TCAGATCAGAACATTGCCAAGCTACTTCCCAATCAACAACACCCCTTCAGGATTTAAANAACCAAGGGGA
TCAGATCAGAACATTGCCAAGCTACTTCCCAATCAACAACACCCCTTCAGGATTTAAANAACCAAGGGGA
150 160 170 180 190 200 210
CACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCGTGTAGGAGAGAGAACTAAAGTTCCTGAAGA
CACTGGATCACCTAGTGTTCACAAGCAGGTACCTTCGTGTAGGAGAGAGAACTAAAGTTCCTGAAGA
150 160 170 180 190 200 210
220 230 240 250 260 270 280
CCTGTGCTTTTACACAGGAAGTTTACTGGGCATCTCTGAGCCTAGGCAATAGCTGTAGGGTCACTTCG
CCTGTGCTTTTACACAGGAAGTTTACTGGGCATCTCTGAGCCTAGGCAATAGCTGTAGGGTCACTTCG
220 230 240 250 260 270 280
290 300 310 320 330 340 350 360
GAGCCATCCCGTTTCCCGCGCCCAAGAGCGGAGATTAAACGGGGAGCTGCGGCCAGAGCTGGGGA
GAGCCATCCCGTTTCCCGCGCCCAAGAGCGGAGATTAAACGGGGAGCTGCGGCCAGAGCTGGGGA
290 300 310 320 330 340 350 360
370 380 390 400 410 420 430
ATGGCCCGGAGCGAGCGGCTTCCTCTGATGCTTTTGCAGACCGCGTCTCGCAGGGCGCTTG
ATGGCCCGGAGCGAGCGGCTTCCTCTGATGCTTTTGCAGACCGCGTCTCGCAGGGCGCTTG
370 380 390 400 410 420 430
440 450 460 470 480 490 500
CTGCGTGAAGTTCGAGGCTTCGGGCACTAGGGCGCGCGGGGTGAAAAATCGAAACTAGCTTTTCT
CTGCGTGAAGTTCGAGGCTTCGGGCACTAGGGCGCGCGGGGTGAAAAATCGAAACTAGCTTTTCT
440 450 460 470 480 490 500
510 520 530 540 550 560 570
TTGCGTTCGGAGTTTGAAGGACCTGCTCAACCTATCCGCAAGCCCTCTCCCTACTTCTG
TTGCGTTCGGAGTTTGAAGGACCTGCTCAACCTATCCGCAAGCCCTCTCCCTACTTCTG
510 520 530 540 550 560 570
580 590 600 610 620 630 640
CGTCCAGACCCCTGAGGAGTGCCTACCACTGAAGTGCAGATAGGGTCCCTCGCCCGCAGAGCTGCCCC
CGTCCAGACCCCTGAGGAGTGCCTACCACTGAAGTGCAGATAGGGTCCCTCGCCCGCAGAGCTGCCCC
580 590 600 610 620 630 640
650 660 670 680 690 700 710 720
TCGCCGGCTGTCCCGGCTGTGGGAGTGACTTTTGAACCCCGCCTCCCTCCCGCACTAGAACTGCTTT
TCGCCGGCTGTCCCGGCTGTGGGAGTGACTTTTGAACCCCGCCTCCCTCCCGCACTAGAACTGCTTT
650 660 670 680 690 700 710 720
730 740 750 760 770 780 790
TAAATAATCTCGTAGTTCCTCACTTCTGAGCTAGCTAAGCCTGGGCTCTTGAACCTGGAAGCTCGGGTTA
TAAATAATCTCGTAGTTCCTCACTTCTGAGCTAGCTAAGCCTGGGCTCTTGAACCTGGAAGCTCGGGTTA
730 740 750 760 770 780 790
800 810 820 830 840 850 860
TTTCCAATGTCACTGTGAGTTTTCCTCCAGTTCATCTCCAAACAGGAGTTCCTCCCTGAGTCTGCCG
TTTCCAATGTCACTGTGAGTTTTCCTCCAGTTCATCTCCAAACAGGAGTTCCTCCCTGAGTCTGCCG
800 810 820 830 840 850 860

870 880 890 900 910 920 930
AGAAGCTGAGCAAAACCCAGCAGGATCCGACGGGTTTCCACCTCAGAACGATGCTTGGCGGTGGG
AGAAGCTGAGCAAAACCCAGCAGGATCCGACGGGTTTCCACCTCAGAACGATGCTTGGCGGTGGG
870 880 890 900 910 920 930
940 950 960 970 980 990 1000
GGCGGAAAGAGTGGGATCTGAATCTTCCACCTCCACCCACTTTTGGTGAGACCTGGGGTGA
GGCGGAAAGAGTGGGATCTGAATCTTCCACCTCCACCCACTTTTGGTGAGACCTGGGGTGA
940 950 960 970 980 990 1000
1010 1020 1030 1040 1050 1060 1070 1080
GGTCTCTAGGGTGGGAGTCCCTGAGAGGCGCTTCTGGGCTTCCACCTCTTGGCAATTTGCTTTT
GGTCTCTAGGGTGGGAGTCCCTGAGAGGCGCTTCTGGGCTTCCACCTCTTGGCAATTTGCTTTT
1010 1020 1030 1040 1050 1060 1070 1080
1090 1100 1110 1120 1130 1140 1150
GCCTGGAATTAAGTATATGTTAGTCTTGAAGCTTGAACCTGAACAAATCTTTTCGGCTAGGCTTAT
GCCTGGAATTAAGTATATGTTAGTCTTGAAGCTTGAACCTGAACAAATCTTTTCGGCTAGGCTTAT
1090 1100 1110 1120 1130 1140 1150
1160 1170 1180 1190 1200 1210 1220
GATTTGCAATGTCTGTATTAAGAGGCTCTCTACAAAGTACTGATTAATGAACATGTAAGCAATGCACT
GATTTGCAATGTCTGTATTAAGAGGCTCTCTACAAAGTACTGATTAATGAACATGTAAGCAATGCACT
1160 1170 1180 1190 1200 1210 1220
1230 1240 1250 1260 1270 1280 1290
CACTTCTAAGTTACATTCATATCTGATTTTATTTGATTTTCACTAGGATAGGAGTAGGACTTAATA
CACTTCTAAGTTACATTCATATCTGATTTTATTTGATTTTCACTAGGATAGGAGTAGGACTTAATA
1230 1240 1250 1260 1270 1280 1290
1300 1310 1320 1330 1340 1350 1360
CGTTTATTTTACTAGAAGTTAAGTGAATTCAGATTAATTAACCTTTTCAGGTACAAAGAACAATAA
CGTTTATTTTACTAGAAGTTAAGTGAATTCAGATTAATTAACCTTTTCAGGTACAAAGAACAATAA
1300 1310 1320 1330 1340 1350 1360
1370 1380 1390 1400 1410 1420 1430 1440
TCTGGTTTCTGATGTTTCAAGTACTACAGCTCTTAATCTTAGTTGACAGTATTTTGGCTATCTCC
TCTGGTTTCTGATGTTTCAAGTACTACAGCTCTTAATCTTAGTTGACAGTATTTTGGCTATCTCC
1370 1380 1390 1400 1410 1420 1430 1440
1450 1460 1470 1480 1490 1500 1510
TGTAGCAGAGTCTCTGGGTACAGCGCGCTCAGCAGCACTTTGAGTTTGGTACTACGTATCC
TGTAGCAGAGTCTCTGGGTACAGCGCGCTCAGCAGCACTTTGAGTTTGGTACTACGTATCC
1450 1460 1470 1480 1490 1500 1510
1520 1530 1540 1550 1560 1570 1580
ACATTTTACATCAGCAAGATGAGGATGCGCGCTCTCTGCAAAATTTATCAATGGTACACTGG
ACATTTTACATCAGCAAGATGAGGATGCGCGCTCTCTGCAAAATTTATTTCAATGGTACACTGG
1520 1530 1540 1550 1560 1570 1580
1590 1600 1610 1620 1630 1640 1650
GCTTTGGTGGCAGAGCTCATGCTCCACTTTCATAGCTATGATTTCTAAACATCAGCTGCAATGAGGTTGA
GCTTTGGTGGCAGAGCTCATGCTCCACTTTCATAGCTATGATTTCTAAACATCAGCTGCAATGAGGTTGA
1590 1600 1610 1620 1630 1640 1650
1660 1670 1680 1690 1700 1710 1720
ATAATAAATTTTATGTTGAGCAAAATTTTCAATGTTTACAAGTGAATGAGTCCAGCCATGTTGCA
ATAATAAATTTTATGTTGAGCAAAATTTTCAATGTTTACAAGTGAATGAGTCCAGCCATGTTGCA
1660 1670 1680 1690 1700 1710 1720
1730 1740 1750 1760 1770 1780 1790 1800

TACCATGGCTAGACACACCTTAAACATTTTCTAGAAATCCACGAGCTTTAGTAGAGTCTCTCTAAATCATCAGTA
3460 3470 3480 3490 3500 3510 3520
3530 3540 3550 3560 3570 3580 3590 3600
TTTGGAAATAGGATCTGGGGCAGTGAGGGGTGGCAGCCACGTGTGGCAGAGAAAGACACAAAGGAAAGAGC
TTGGAAATAGGATCTGGGGCAGTGAGGGGTGGCAGCCACGTGTGGCAGAGAAAGACACAAAGGAAAGAGC
3530 3540 3550 3560 3570 3580 3590 3600
ACCCAGGACTGTCATATGGAAGAAAGACAGAGACTGCAACTCACCCCTTCACAAAATGAGGACGACACAGCT
ACCCAGGACTGTCATATGGAAGAAAGACAGAGACTGCAACTCACCCCTTCACAAAATGAGGACGACACAGCT
3610 3620 3630 3640 3650 3660 3670
3680 3690 3700 3710 3720 3730 3740
GATGGTATAGTGTGATGCAAGGTGTGGAGGCTCAACATCTCGTCCCTCTCTACTACACATGGTTAAGGCC
GATGGTATAGTGTGATGCAAGGTGTGGAGGCTCAACATCTCGTCCCTCTCTACTACACATGGTTAAGGCC
3680 3690 3700 3710 3720 3730 3740
3750 3760 3770 3780 3790 3800 3810
TGTTGCTCTGTCTCAGGTTACACTCTCTGCACTACCTCTTCATGGGTGCTCAGACGAGACCTTGGTCT
3750 3760 3770 3780 3790 3800 3810
3820 3830 3840 3850 3860 3870 3880
TTCCCTTGTGTAAGCTTTGGGCTACGTGTGATGACGAGCTGTTTCGTGTCTATGATNATGAGAGTGCCTGTGT
TTCCCTTGTGTAAGCTTTGGGCTACGTGTGATGACGAGCTGTTTCGTGTCTATGATGATGAGAGTGCCTGTGT
3820 3830 3840 3850 3860 3870 3880
3890 3900 3910 3920 3930 3940 3950 3960
GGAGCCCGAAGCTCCATGGTTTCCAGTGAATTTCAAGCCAGATGTGGCTCAGCTGAGTCAGAGTCTGAA
GGAGCCCGAAGCTCCATGGTTTCCAGTGAATTTCAAGCCAGATGTGGCTCAGCTGAGTCAGAGTCTGAA
3890 3900 3910 3920 3930 3940 3950 3960
3970 3980 3990 4000 4010 4020 4030
AGGGTGGATACATCTTCACTGTTCAGTCTTGAGCTATTATGGAATTCACAAACACAGCAAGGATGTG
AGGGTGGATACATCTTCACTGTTCAGTCTTGAGCTATTATGGAATTCACAAACACAGCAAGGATGTG
3970 3980 3990 4000 4010 4020 4030
4040 4050 4060 4070 4080 4090 4100
GAGAGGGGCTCACCTTCTGAGGTGTGACAGCTTTTCATCTTTTCATGATCTTTGAGGAAACAGCTGG
GAGAGGGGCTCACCTTCTGAGGTGTGACAGCTTTTCATCTTTTCATGATCTTTGAGGAAACAGCTGG
4040 4050 4060 4070 4080 4090 4100
4110 4120 4130 4140 4150 4160 4170
AAGTCTCAGGCTTGTGGAGCAGGGAAGGGAAGGAATTTGCTTCCTGAGATCATTTTGGTCCCTGGGGAT
AAGTCTCAGGCTTGTGGAGCAGGGAAGGGAAGGAATTTGCTTCCTGAGATCATTTTGGTCCCTGGGGAT
4110 4120 4130 4140 4150 4160 4170
AAGTCTCAGGCTTGTGGAGCAGGGAAGGGAAGGAATTTGCTTCCTGAGATCATTTTGGTCCCTGGGGAT
AAGTCTCAGGCTTGTGGAGCAGGGAAGGGAAGGAATTTGCTTCCTGAGATCATTTTGGTCCCTGGGGAT
4110 4120 4130 4140 4150 4160 4170
4180 4190 4200 4210 4220 4230 4240
GGTGGAAATAGGACCTATTCCCTTGGTGCAGTTAACAGGCTGGGGATTTTCCAGAGTCCCAACCCCTG
GGTGGAAATAGGACCTATTCCCTTGGTGCAGTTAACAGGCTGGGGATTTTCCAGAGTCCCAACCCCTG
4180 4190 4200 4210 4220 4230 4240
4250 4260 4270 4280 4290 4300 4310 4320
CAGGTATCTGGGCTGTGAAATGCAAGAGACACAGTACCCAGGGCTACTGGAAGTACGGGTATGATGGG
CAGGTATCTGGGCTGTGAAATGCAAGAGACACAGTACCCAGGGCTACTGGAAGTACGGGTATGATGGG
4250 4260 4270 4280 4290 4300 4310 4320
4330 4340 4350 4360 4370 4380 4390
CAGGACCCACTTGAATTTCTGCCCTGACACACTGATTTGGAGAGCAGACAGAACCCAGGGCTGSCCCCAACG
CAGGACCCACTTGAATTTCTGCCCTGACACACTGATTTGGAGAGCAGACAGAACCCAGGGCTGSCCCCAACG
4330 4340 4350 4360 4370 4380 4390
CAGGACCCACTTGAATTTCTGCCCTGACACACTGATTTGGAGAGCAGACAGAACCCAGGGCTGSCCCCAACG
CAGGACCCACTTGAATTTCTGCCCTGACACACTGATTTGGAGAGCAGACAGAACCCAGGGCTGSCCCCAACG

CTGAGTGGGAAAGGACAAAGATTCTGGGCCAGGCAGAACAGGGCCCTACCTGGAGAGGGACTGCCCTGCACAG
CTGAGTGGGAAAGGACAAAGATTCTGGGCCAGGCAGAACAGGGCCCTACCTGGAGAGGGACTGCCCTGCACAG
4400 4410 4420 4430 4440 4450 4460
4470 4480 4490 4500 4510 4520 4530
CTGAGCAGTGTCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGGTATGTTGGAACACACTTCTGCCTCC
CTGAGCAGTGTCTGGAGCTGGGGAGAGGTGTTTGGACCAACAAGGTATGTTGGAACACACTTCTGCCTCC
4470 4480 4490 4500 4510 4520 4530
4540 4550 4560 4570 4580 4590 4600
TATACTTAGTGGCAGAGTGGAGAGGTTCAGGGCAGCGAATCCCTGGTTGGAGTTTCAGAGTGGCTGAG
TATACTTAGTGGCAGAGTGGAGAGGTTCAGGGCAGCGAATCCCTGGTTGGAGTTTCAGAGTGGCTGAG
4540 4550 4560 4570 4580 4590 4600
4610 4620 4630 4640 4650 4660 4670 4680
GCTGTGTCTCTCCAAATCTCTGGAAGGACTTTCATCTCTAGAGTCTCTACCTTATAATTGAGATGTA
GCTGTGTCTCTCCAAATCTCTGGAAGGACTTTCATCTCTAGAGTCTCTACCTTATAATTGAGATGTA
4610 4620 4630 4640 4650 4660 4670 4680
4690 4700 4710 4720 4730 4740 4750
TGAGACGCCACAAAGTCATGGGTTTAATTTCTTCTCCATGCATATGGCTCAAGGAAAGTGTCTATGGCC
TGAGACGCCACAAAGTCATGGGTTTAATTTCTTCTCCATGCATATGGCTCAAGGAAAGTGTCTATGGCC
4690 4700 4710 4720 4730 4740 4750
4760 4770 4780 4790 4800 4810 4820
CTTGCTTTTATTAAACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGAAATGTCAAGCCGGG
CTTGCTTTTATTAAACCAATAATCTTTTGTATATTATACCTGTTAAAAATTCAGAAATGTCAAGCCGGG
4760 4770 4780 4790 4800 4810 4820
4830 4840 4850 4860 4870 4880 4890
CACGGTGGCTCACCCCTGTAATCCAGCACTTTTGGAGGCGGAGGGGTGTCACAAAGTCAAGGATTTGA
CACGGTGGCTCACCCCTGTAATCCAGCACTTTTGGAGGCGGAGGGGTGTCACAAAGTCAAGGATTTGA
4830 4840 4850 4860 4870 4880 4890
4900 4910 4920 4930 4940 4950 4960
GACGAGCTGACCAACATGTTGGAACCCGTCCTTAAAAAATACAAAAATAGCTGTGCACAGTCATGCGCA
GACGAGCTGACCAACATGTTGGAACCCGTCCTTAAAAAATACAAAAATAGCTGTGCACAGTCATGCGCA
4900 4910 4920 4930 4940 4950 4960
4970 4980 4990 5000 5010 5020 5030 5040
CCTGTAGTCCGAGCTAATTTGAAGGCTGAGCGAGGATCGCTTGAACCTTGAACCTGGGAAGGTTGCACCTGA
CCTGTAGTCCGAGCTAATTTGAAGGCTGAGCGAGGATCGCTTGAACCTTGAACCTGGGAAGGTTGCACCTGA
4970 4980 4990 5000 5010 5020 5030 5040
5050 5060 5070 5080 5090 5100 5110
GCCAGATCGGGCCACTGCATCCAGCTTGGAGGAGTGAAGTCCATCTTAAAAAATACAAAAA
GCCAGATCGGGCCACTGCATCCAGCTTGGAGGAGTGAAGTCCATCTTAAAAAATACAAAAA
5050 5060 5070 5080 5090 5100 5110
5120 5130 5140 5150 5160 5170 5180
AAAAAGAAATTCAGAGATCTCAGCTATCATATGAATACGAGCAAAATATCAATGAGGCCCTATCAG
AAAAAGAAATTCAGAGATCTCAGCTATCATATGAATACGAGCAAAATATCAATGAGGCCCTATCAG
5120 5130 5140 5150 5160 5170 5180
5190 5200 5210 5220 5230 5240 5250
AGTAGAAGAAATTCAGAGTAAAAAGTTTCTTCTATAGAACATAGCAATATCACTGAAGCTACCTCTCTTA
AGTAGAAGAAATTCAGAGTAAAAAGTTTCTTCTATAGAACATAGCAATATCACTGAAGCTACCTCTCTTA
5190 5200 5210 5220 5230 5240 5250

5260 5270 5280 5290 5300 5310 5320
CAGTCGCTCTTATATAAATGCCCTCCTAGGTTGACCCAGGTGAAGTGAACATCTGTATTCATCATTTTT
CAGTCGCTCTTATATAAATGCCCTCCTAGGTTGACCCAGGTGAAGTGAACATCTGTATTCATCATTTTT
5260 5270 5280 5290 5300 5310 5320
5330 5340 5350 5360 5370 5380 5390 5400
CAATGCACATAAAGGCCAATTTATCTATCAGAACAAAGACATGGGTAAACAGATATATATTTACATGTG
|||||
CAATGCACATAAAGGCCAATTTATCTATCAGAACAAAGACATGGGTAAACAGATATATATTTACATGTG
5330 5340 5350 5360 5370 5380 5390 5400
5410 5420 5430 5440 5450 5460 5470
AGGAGAACAGCTGATCTGACTGCTCTCCAGTGACACTGTGTAGAGTCCCAATCTTAGGACACAAAATGT
|||||
AGGAGAACAGCTGATCTGACTGCTCTCCAGTGACACTGTGTAGAGTCCCAATCTTAGGACACAAAATGT
5410 5420 5430 5440 5450 5460 5470
5480 5490 5500 5510 5520 5530 5540
GTCTCTCCGTAGCTGTTTTTCTGAAAGGTTATTTCTCTCCCAACTATAGAGGAAGTGAAGTT
|||||
GTCTCTCCGTAGCTGTTTTTCTGAAAGGTTATTTCTCTCCCAACTATAGAGGAAGTGAAGTT
5480 5490 5500 5510 5520 5530 5540
5550 5560 5570 5580 5590 5600 5610
CCAGTCTTCTCGCAAGGTAACAGATCCCTCTCCCTCATCTTCCTTCCTCGTCAAGTGCCTCTTTG
|||||
CCAGTCTTCTCGCAAGGTAACAGATCCCTCTCCCTCATCTTCCTTCCTCGTCAAGTGCCTCTTTG
5550 5560 5570 5580 5590 5600 5610
5620 5630 5640 5650 5660 5670 5680
GTGAAGGTGACACATCATGACCTCTTCAGTGACCCTCTACGGTTCGGGGCCTTGAACCTACTACCCCGAG
|||||
GTGAAGGTGACACATCATGACCTCTTCAGTGACCCTCTACGGTTCGGGGCCTTGAACCTACTACCCCGAG
5620 5630 5640 5650 5660 5670 5680
5690 5700 5710 5720 5730 5740 5750 5760
AATCATCAGCTAAGTGCCTGAAGGATGAAGCAGCAATGGATGTCAGAGGTTTGAACCTTAAAGACGTATG
5690 5700 5710 5720 5730 5740 5750 5760
5770 5780 5790 5800 5810 5820 5830
CCCAATGGGATGGACCTACAGGGCTGGATACCTTGGCTGTACCCCTGGGGAGAGCAGAGATATAG
|||||
CCCAATGGGATGGACCTACAGGGCTGGATACCTTGGCTGTACCCCTGGGGAGAGCAGAGATATAG
5770 5780 5790 5800 5810 5820 5830
5840 5850 5860 5870 5880 5890 5900
TNCCAGGTGGAGCACCAGGCTGGATCAGCCCTCATTTGGATCTGGGTATGTGACTGATGAGACCCAGG
|||||
TNCCAGGTGGAGCACCAGGCTGGATCAGCCCTCATTTGGATCTGGGTATGTGACTGATGAGACCCAGG
5840 5850 5860 5870 5880 5890 5900
5910 5920 5930 5940 5950 5960 5970
AGCTGAGAAAATCTATTGGGGTTGAGAGGAGTCCCTGAGGAGTAAATATGGCAGTGAGATGAGATCTGC
|||||
AGCTGAGAAAATCTATTGGGGTTGAGAGGAGTCCCTGAGGAGTAAATATGGCAGTGAGATGAGATCTGC
5910 5920 5930 5940 5950 5960 5970
5980 5990 6000 6010 6020 6030 6040
TCTTTGTAGGGGTGGGCTGAGGTTGCAATCAAGGCTTAACTTGTCTTTCTGTTTACAGCCCTCAC
|||||
TCTTTGTAGGGGTGGGCTGAGGTTGCAATCAAGGCTTAACTTGTCTTTTCTGTTTACAGCCCTCAC
5980 5990 6000 6010 6020 6030 6040
6050 6060 6070 6080 6090 6100 6110 6120
CGCTGGCACCCCTAGTCATTTGAGTGCATGCAATGCTTTTGTGCTCATCTTGTTCATTGGAATTT
|||||
CGCTGGCACCCCTAGTCATTTGAGTGCATGCAATGCTTTTGTGCTCATCTTGTTCATTGGAATTT
6050 6060 6070 6080 6090 6100 6110 6120

6130 6140 6150 6160 6170 6180 6190
TGTTTCATAAATTAAGGAAGAGGAGGTTCAAGTGAAGTGAAGCAAGGGGAAGTCTCTTAGTACCTCTGC
|||||
TGTTTCATAAATTAAGGAAGAGGAGGTTCAAGTGAAGTGAAGCAAGGGGAAGTCTCTTAGTACCTCTGC
6130 6140 6150 6160 6170 6180 6190
6200 6210 6220 6230 6240 6250 6260
CCCAGGCACAGTGGGAAGAGGAGGAGGATCTGGCATCCATGGGAAGCATTTTCTCATTTATATCT
|||||
CCCAGGCACAGTGGGAAGAGGAGGAGGATCTGGCATCCATGGGAAGCATTTTCTCATTTATATCT
6200 6210 6220 6230 6240 6250 6260
6270 6280 6290 6300 6310 6320 6330
TTGGGGACACAGCAGCTCCCTGGGAGACAGAAAATATGTTCTCCCAAGATGAAGTCTCTAATTTCAAC
|||||
TTGGGGACACAGCAGCTCCCTGGGAGACAGAAAATATGTTCTCCCAAGATGAAGTCTCTAATTTCAAC
6270 6280 6290 6300 6310 6320 6330
6340 6350 6360 6370 6380 6390 6400
AAACATCTTCAGAGCACCTACTATTTTGCAGAGCTGTTTAAAGTAGTACAGGGGCTTTGAGGTTGAGAAGT
|||||
AAACATCTTCAGAGCACCTACTATTTTGCAGAGCTGTTTAAAGTAGTACAGGGGCTTTGAGGTTGAGAAGT
6340 6350 6360 6370 6380 6390 6400
6410 6420 6430 6440 6450 6460 6470 6480
CACTGTGGCTTATTTCTCAGAACCCCAAAATCTGGTAGGGAATGAAATTTGATAGCAAGTAAATGTTAAAGAAG
|||||
CACTGTGGCTTATTTCTCAGAACCCCAAAATCTGGTAGGGAATGAAATTTGATAGCAAGTAAATGTTAAAGAAG
6410 6420 6430 6440 6450 6460 6470 6480
6490 6500 6510 6520 6530 6540 6550
ACCCCATGAGGTCCTTAAAGCAGCAGGAAGCAATGCTTAGGGTGTCAAAGAAAGAAATGATCACAATTCAGC
|||||
ACCCCATGAGGTCCTTAAAGCAGCAGGAAGCAATGCTTAGGGTGTCAAAGAAAGAAATGATCACAATTCAGC
6490 6500 6510 6520 6530 6540 6550
6560 6570 6580 6590 6600 6610 6620
TGGGATCAAGATAGCTTCTGGATCTTGAAGGAGAGCTGGATTCCATTAGGTAGGTTGAAGATGATGGG
6560 6570 6580 6590 6600 6610 6620
6630 6640 6650 6660 6670 6680 6690
AGGTCTACACAGCAGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATCTGGGAGATTAGAATAATTAC
|||||
AGGTCTACACAGCAGGAGCAACCATGCCAAGTAGGAGAGTATAAGGCATCTGGGAGATTAGAATAATTAC
6630 6640 6650 6660 6670 6680 6690
6700 6710 6720 6730 6740 6750 6760
TGTTACCTTAACCTGAGTTTGGTAGCTATCTACTACCAATATGCAATTTACCCCTGAACATCTGTGGT
|||||
TGTTACCTTAACCTGAGTTTGGTAGCTATCTACTACCAATATGCAATTTACCCCTGAACATCTGTGGT
6700 6710 6720 6730 6740 6750 6760
6770 6780 6790 6800 6810 6820 6830 6840
GTAGGAAAGAGAGTAATCAGAAAGAACCCAGCTCATACAGAGTCCCAAGGCTTTTGGGATATTTGGGTATGA
|||||
GTAGGAAAGAGAGTAATCAGAAAGAACCCAGCTCATACAGAGTCCCAAGGCTTTTGGGATATTTGGGTATGA
6770 6780 6790 6800 6810 6820 6830 6840
6850 6860 6870 6880 6890 6900 6910
TCACTGGGTGTCTATTGAAGGATCCCTAAGAAAGGAGGACCAAGTCTCCCTTATATGTTGAATGTGTGTTA
|||||
TCACTGGGTGTCTATTGAAGGATCCCTAAGAAAGGAGGACCAAGTCTCCCTTATATGTTGAATGTGTGTTA
6850 6860 6870 6880 6890 6900 6910
6920 6930 6940 6950 6960 6970 6980
AGAAGTTAGATGAGAGGTGAGGAGCCAGTTAGAAGCCCAATAGCAATTCAGATGAGAGATAATGTTCT
|||||
AGAAGTTAGATGAGAGGTGAGGAGCCAGTTAGAAGCCCAATAGCAATTCAGATGAGAGATAATGTTCT
6920 6930 6940 6950 6960 6970 6980
6990 7000 7010 7020 7030 7040 7050

TCCTGGAATGCTGACTCCCTTGCTCCTCTGTTCTCTCTCTGTCATTCATTTCTTTGGACCCCTACGCAAGGAC
8720 8730 8740 8750 8760 8770 8780
8790 8800 8810 8820 8830 8840 8850
TCGTAATTTGGTGGGACAGCTAGTGGCCCTGCTGGGCTTCACACACGGTGTCTCCCTAGGCCAGTGCCTCG
TCGTAATTTGGTGGGACAGCTAGTGGCCCTGCTGGGCTTCACACACGGTGTCTCCCTAGGCCAGTGCCTCG
8790 8800 8810 8820 8830 8840 8850
8860 8870 8880 8890 8900 8910 8920
GAGTCAGAACTCTGCTGGTATTTCCCTCAATGAAGTGCAGTAAAGCTCTCTCATTTTGAGATGGTAAATGA
GAGTCAGAACTCTGCTGGTATTTCCCTCAATGAAGTGCAGTAAAGCTCTCTCATTTTGAGATGGTAAATGA
8860 8870 8880 8890 8900 8910 8920
8930 8940 8950 8960 8970 8980 8990 9000
AGCCACCAAGTGGCTTAGAGATGCCAGGTCCCTCCATGGAGCCAGTCCGGTGCACATTAATAAA
AGCCACCAAGTGGCTTAGAGATGCCAGGTCCCTCCATGGAGCCAGTCCGGTGCACATTAATAAA
8930 8940 8950 8960 8970 8980 8990 9000
9010 9020 9030 9040 9050 9060 9070
AAAACTAACCCAGGACATTCAGGAATTCCTAGATTCCTGGAAATCAGTTCACCATGTTCAAAGAGTCTTT
AAAACTAACCCAGGACATTCAGGAATTCCTAGATTCCTGGAAATCAGTTCACCATGTTCAAAGAGTCTTT
9010 9020 9030 9040 9050 9060 9070
9080 9090 9100 9110 9120 9130 9140
TTTTTTTTTGGAGCTCTATTGCCAGGCTGGAGTGCATGCGCTCACTGTAAAGCTGTGCT
TTTTTTTTTGGAGCTCTATTGCCAGGCTGGAGTGCATGCGCTCACTGTAAAGCTGTGCT
9080 9090 9100 9110 9120 9130 9140
9150 9160 9170 9180 9190 9200 9210
CCGAGGTTCAAGCCGATTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGCGTGCACCCAGTGCCTG
CCGAGGTTCAAGCCGATTCCTGTCTCAGCCTCCCAAGTAGCTGGGATTACAGCGTGCACCCAGTGCCTG
9150 9160 9170 9180 9190 9200 9210
9220 9230 9240 9250 9260 9270 9280
GCTAAATTTGCTATTTTAGTAGACAGGTTTCACCATGTTGGCCAGGCTGGTTCGCAACTCTCCTGACC
GCTAAATTTGCTATTTTAGTAGACAGGTTTCACCATGTTGGCCAGGCTGGTTCGCAACTCTCCTGACC
9220 9230 9240 9250 9260 9270 9280
9290 9300 9310 9320 9330 9340 9350 9360
TCGTATCCGCTCGGCTCCCAAGTGTGAGATTACAGGTTGAGCCACCTGCGCCAGCCGTCAA
TCGTATCCGCTCGGCTCCCAAGTGTGAGATTACAGGTTGAGCCACCTGCGCCAGCCGTCAA
9290 9300 9310 9320 9330 9340 9350 9360
9370 9380 9390 9400 9410 9420 9430
AGAGTCTTAATATATATATCCAGATGGCATGTTTACTTTATGTTACTACATGCACTGGCTGCATAAATG
AGAGTCTTAATATATATATCCAGATGGCATGTTTACTTTATGTTACTACATGCACTGGCTGCATAAATG
9370 9380 9390 9400 9410 9420 9430
9440 9450 9460 9470 9480 9490 9500
TGTACAAAGCATTCGCTTGAAGGCGAGTGTTCAGGATACCATATACAGTTCAGAGATTTCTTCTTAG
TGTACAAAGCATTCGCTTGAAGGCGAGTGTTCAGGATACCATATACAGTTCAGAGATTTCTTCTTAG
9440 9450 9460 9470 9480 9490 9500
9510 9520 9530 9540 9550 9560 9570
GCATTAATTTAGCAAGATATCTCATCTCTCTTTTAAACCATTTCTTTTGGTGTAGAAAAGTTA
GCATTAATTTAGCAAGATATCTCATCTCTCTTTTAAACCATTTCTTTTGGTGTAGAAAAGTTA
9510 9520 9530 9540 9550 9560 9570
9580 9590 9600 9610 9620 9630 9640
TGTAGAAAAAGTAAATGTGATTTACGCTCATTTGAGAAAAGCTATAAATGATACAAATTAAGCTGTAT
TGTAGAAAAAGTAAATGTGATTTACGCTCATTTGAGAAAAGCTATAAATGATACAAATTAAGCTGTAT

9580 9590 9600 9610 9620 9630 9640
9650 9660 9670 9680 9690 9700 9710 9720
TTAATAGCCAGTGA AAAAATTAATTAACAACCTTGCTCTATTACCTGTTAGTATTATTTGTCATTTAAATAATGCA
TTAATAGCCAGTGA AAAAATTAATTAACAACCTTGCTCTATTACCTGTTAGTATTATTTGTCATTTAAATAATGCA
9650 9660 9670 9680 9690 9700 9710 9720
TATACCTTAATAAATGATATATGTTATGTTACTGCTATGATGATTTTATGAAGTCTTGTCACTCTCTGTATA
TATACCTTAATAAATGATATATGTTATGTTACTGCTATGATGATTTTATGAAGTCTTGTCACTCTCTGTATA
9730 9740 9750 9760 9770 9780 9790 9800
TACTTAATCCCTTTGTCATTTTGGAGACATTTATTTGGCTCTAAATTTCTTTACATTTTGTCTTACGGAATA
TACTTAATCCCTTTGTCATTTTGGAGACATTTATTTTGGCTCTAAATTTCTTTACATTTTGTCTTACGGAATA
9800 9810 9820 9830 9840 9850 9860 9870
TTTTCATTCACCTGCTGGTAGCCGAATTAATTCGTTTCTTCACTCTAGGACATTTCTCTCTAAGTGTGAAG
TTTTCATTCACCTGCTGGTAGCCGAATTAATTCGTTTCTTCACTCTAGGACATTTCTCTCTAAGTGTGAAG
9870 9880 9890 9900 9910 9920 9930 9940
ACATTTGGTATTTTACCAGCAACCATTTCTGAAAGCATATGACAAATTTATTTCTCTTAAATCTTACTAT
ACATTTGGTATTTTACCAGCAACCATTTCTGAAAGCATATGACAAATTTATTTCTCTTAAATCTTACTAT
9940 9950 9960 9970 9980 9990 10000 10010
ACTGAAAGCAGACTGCTATAAGGCTTCACTTACTCTTCACTCTACCTCATAAGGAATATGTTACAAATTTATTT
ACTGAAAGCAGACTGCTATAAGGCTTCACTTACTCTTCACTCTACCTCATAAGGAATATGTTACAAATTTATTT
10010 10020 10030 10040 10050 10060 10070 10080
AGTAAAGCATTTGTTTATATATGTTTATTTTCACTGGCTGAGATTTCAAGAACACCCAGCTTCACCA
AGTAAAGCATTTGTTTATATGTTTATTTTCACTGGCTGAGATTTCAAGAACACCCAGCTTCACCA
10090 10100 10110 10120 10130 10140 10150
10160 10170 10180 10190 10200 10210 10220
GTAACACATTTCACTAAACACATTTACTAAACATCAGCAACTGTGGCTGTTAAATTTTAAATAGAAATTT
GTAACACATTTCACTAAACACATTTACTAAACATCAGCAACTGTGGCTGTTAAATTTTAAATAGAAATTT
10160 10170 10180 10190 10200 10210 10220
10230 10240 10250 10260 10270 10280 10290
AAGTCTCATTTTCTTTCGGTGTGTTTAAAGCTTAAATTTTCTGGCTTATTCATPAATTTCTTAAGGTCAAC
AAGTCTCATTTTCTTTCGGTGTGTTTAAAGCTTAAATTTTCTGGCTTATTCATPAATTTCTTAAGGTCAAC
10230 10240 10250 10260 10270 10280 10290
10300 10310 10320 10330 10340 10350 10360
TACATTTGAAAATCAAGACCTGCAATTTAAATTTCTTATTCACCTCTGGCAAAACCATTTCAAAAACCATGG
TACATTTGAAAATCAAGACCTGCAATTTAAATTTCTTATTCACCTCTGGCAAAACCATTTCAAAAACCATGG
10300 10310 10320 10330 10340 10350 10360
10370 10380 10390 10400 10410 10420 10430 10440
TAGTAAGAGAGGCTGACACCTGTGGCCATAGTAATGTACCACGGTGGTGGGAGACAGAGATGCGAG
TAGTAAGAGAGGCTGACACCTGTGGCCATAGTAATGTACCACGGTGGTGGGAGACAGAGATGCGAG
10370 10380 10390 10400 10410 10420 10430 10440
CGCTGAGGCTTTCTCTGAAGGTAAAGGAATAAAGATGGTGGAGGCGCTGCACCTGAAATCACTTGTAGA
CGCTGAGGCTTTCTCTGAAGGTAAAGGAATAAAGATGGTGGAGGCGCTGCACCTGAAATCACTTGTAGA
10450 10460 10470 10480 10490 10500 10510


```
ACTGAAACGACACTGCTATAGGCTTCACTTACTTCTTACCTCATAGGAATATGTTACAAATTAATTTATT
10010 10020 10030 10040 10050 10060 10070 10080
AGGTAAAGCATTTGTTTTATATGTTTTATTTTACCTGGGCTGAGATTTCAAGAAACACCCAGCTCTTCACA
10090 10100 10110 10120 10130 10140 10150
AGGTAAAGCATTTGTTTTATATGTTTTATTTTACCTGGGCTGAGATTTCAAGAAACACCCAGCTCTTCACA
10090 10100 10110 10120 10130 10140 10150
GTAACACATTTTCACTTAACACATTTTACTTAAACATCAGCAACTGTGGCTGTAAATTTTTTAAATGAAATTTT
10160 10170 10180 10190 10200 10210 10220
GTAACACATTTTCACTTAACACATTTTACTTAAACATCAGCAACTGTGGCTGTAAATTTTTTAAATGAAATTTT
10160 10170 10180 10190 10200 10210 10220
10230 10240 10250 10260 10270 10280 10290
AAGTCCTCATTTTCTTTCGGTGTGTTTTAAAGCTTAATTTTCTGGCTTTATTCATAAATTTCTTAAGGTCAAC
10230 10240 10250 10260 10270 10280 10290
10300 10310 10320 10330 10340 10350 10360
TACATTTGAAAAATCAAAACACCTGCATTTTAAATTTTATTCACCTTGGCAAAACCATTCACAAACCATG
10300 10310 10320 10330 10340 10350 10360
TAGTAAGAGAGGGTGACACCTGGTGGCCATAGGTAAATGTACCACGCTGTGCGGTGACCAGAGATGTCAG
10370 10380 10390 10400 10410 10420 10430 10440
TAGTAAGAGAGAGGGTGACACCTGGTGGCCATAGGTAAATGTACCACGCTGTGCGGTGACCAGAGATGTCAG
10370 10380 10390 10400 10410 10420 10430 10440
10450 10460 10470 10480 10490 10500 10510
CGCTGAGGGTTTTCTTGAAGGTAAGGAATAAGAAATGGTGGAGGGGCTGCACTGGAAATCACTTGTAGA
10450 10460 10470 10480 10490 10500 10510
CGCTGAGGGTTTTCTTGAAGGTAAGGAATAAGAAATGGTGGAGGGGCTGCACTGGAAATCACTTGTAGA
10450 10460 10470 10480 10490 10500 10510
10520 10530 10540 10550 10560 10570 10580
GAAAGCCCTGAAAATTTGAGAAACAACAAGAACTACTTACCAGCTATTGGAATTCGTTGGAACTCACAG
10520 10530 10540 10550 10560 10570 10580
GAAAGCCCTGAAAATTTGAGAAACAACAAGAACTACTTACCAGCTATTGGAATTCGTTGGAACTCACAG
10520 10530 10540 10550 10560 10570 10580
10590 10600 10610 10620 10630 10640 10650
GCCATTGCTGAGTGCTGAACTGGGAACAACAAGAACTACTTACCAGCTATTGGAATTCGTTGGAACTCAA
10590 10600 10610 10620 10630 10640 10650
GCCATTGCTGAGTGCTGAACTGGGAACAACAAGAACTACTTACCAGCTATTGGAATTCGTTGGAACTCAA
10590 10600 10610 10620 10630 10640 10650
10660 10670 10680 10690 10700 10710 10720
GTACAGCAGGTGATTTGAGGACTGCTGAGAGGTACAGGCCAAATTTCTATGTTGTTATTAATAATGTGTCATC
10660 10670 10680 10690 10700 10710 10720
GTACAGCAGGTGATTTGAGGACTGCTGAGAGGTACAGGCCAAATTTCTATGTTGTTATTAATAATGTGTCATC
10660 10670 10680 10690 10700 10710 10720
10730 10740 10750 10760 10770 10780 10790 10800
TTATAATGCTGATGTTTATTAATAATTTTCAAACTTTCACAACTCACACATTTTAAACAAACACTGTCTC
10730 10740 10750 10760 10770 10780 10790 10800
TTATAATGCTGATGTTTATTAATAATTTTCAAACTTTCACAACTCACACATTTTAAACAAACACTGTCTC
10730 10740 10750 10760 10770 10780 10790 10800
10810 10820 X
TAAATCCCAAAATTTTTCATAAAC
10810 10820 X
TAAATCCCAAAATTTTTCATAAAC
10810 10820 X
```

```
Initial Score = 10825 Optimized Score = 10825 Significance = 4.02
Residue Identity = 99% Matches = 10823 Mismatches = 2
Gaps = 0 Conservative Substitutions
X 10 20 30 40 50 60 70
TCTAAGGTTGAGATAAAATTTTAAATGATGATTTGAAATTCATAAATATTTAAATATCTAAGT
10 20 30 40 50 60 70
TCTAAGGTTGAGATAAAATTTTAAATGATGATTTGAAATTCATAAATATTTAAATATCTAAGT
10 20 30 40 50 60 70
TCTAAGGTTGAGATAAAATTTTAAATGATGATTTGAAATTCATAAATATTTAAATATCTAAGT
10 20 30 40 50 60 70
TCAGATCAGAACATTTGGAGCTACTTTTCCCAATCAACAAACACCCTTCAGGATTTAAACCAAGGGGA
80 90 100 110 120 130 140
TCAGATCAGAACATTTGGAGCTACTTTTCCCAATCAACAAACACCCTTCAGGATTTAAACCAAGGGGA
80 90 100 110 120 130 140
TCAGATCAGAACATTTGGAGCTACTTTTCCCAATCAACAAACACCCTTCAGGATTTAAACCAAGGGGA
80 90 100 110 120 130 140
150 160 170 180 190 200 210
CACTGGATCACCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGAGAGAGAGAACTAAAGTTCTTGAAGA
150 160 170 180 190 200 210
CACTGGATCACCTAGTGTTCACAAGCAGGTACTTCTGCTGTAGAGAGAGAGAACTAAAGTTCTTGAAGA
150 160 170 180 190 200 210
220 230 240 250 260 270 280
CCTGTGCTTTTACCAGGAGGATTTTACTGGCATCTCTGAGCTAGGCAATAGCTGTAGGCTGACTTCTG
220 230 240 250 260 270 280
CCTGTGCTTTTACCAGGAGGATTTTACTGGCATCTCTGAGCTAGGCAATAGCTGTAGGCTGACTTCTG
220 230 240 250 260 270 280
290 300 310 320 330 340 350 360
GAGCATCCCGTTTCCCGCCCCCAAAAGAGCGAGATTTAACGGGACGCTGCGGCAGAGCTGGGGAA
290 300 310 320 330 340 350 360
GAGCATCCCGTTTCCCGCCCCCAAAAGAGCGAGATTTAACGGGACGCTGCGGCAGAGCTGGGGAA
290 300 310 320 330 340 350 360
ATGGGCCCGCAGCAGCGCGGCTTCTCCTCTGATGCTTTTGCAGAGCGGGTCTGCAGGGGCGCTTG
370 380 390 400 410 420 430
ATGGGCCCGCAGCAGCGCGGCTTCTCCTCTGATGCTTTTGCAGAGCGGGTCTGCAGGGGCGCTTG
370 380 390 400 410 420 430
CTGGCTGAGTCCGAGGCTGCGGGCGAACTAGGGCGCGGGGGTGGAAAAATCGAAATAGCTTTTCT
440 450 460 470 480 490 500
CTGGCTGAGTCCGAGGCTGCGGGCGAACTAGGGCGCGGGGGTGGAAAAATCGAAATAGCTTTTCT
440 450 460 470 480 490 500
510 520 530 540 550 560 570
TTGGCCTTGGGAGTTTGCTAACTTTGGAGGACCTGCTCAACCTATCGCAAGCCCTCTCCCTACTTTCTG
510 520 530 540 550 560 570
TTGGCCTTGGGAGTTTGCTAACTTTGGAGGACCTGCTCAACCTATCGCAAGCCCTCTCCCTACTTTCTG
510 520 530 540 550 560 570
580 590 600 610 620 630 640
CGTCAGACCCCGTGGAGGAGTGCCCTACCTGAACTGCAGATAGGGTCCCTCGCCAGGACCTGCCCC
580 590 600 610 620 630 640
CGTCAGACCCCGTGGAGGAGTGCCCTACCTGAACTGCAGATAGGGTCCCTCGCCAGGACCTGCCCC
580 590 600 610 620 630 640
650 660 670 680 690 700 710 720
TCGCCGGCTGTCGCCGGCTCTGCGGAGTGACTTTTGAACCGCCCACTCCCTTCCCACTAGATGCTTT
650 660 670 680 690 700 710 720
TCGCCGGCTGTCGCCGGCTCTGCGGAGTGACTTTTGAACCGCCCACTCCCTTCCCACTAGATGCTTT
650 660 670 680 690 700 710 720
TAAATAAATCTCGTAGTTCCTACCTTGAAGCTGGGGCTTGAACCTGGAACTCGGGTTA
730 740 750 760 770 780 790
TAAATAAATCTCGTAGTTCCTACCTTGAAGCTGGGGCTTGAACCTGGAACTCGGGTTA
730 740 750 760 770 780 790
800 810 820 830 840 850 860
TTTCCAATGCTCAGCTGTGAGTTTTTCCCGAGTCTCTCCAAACAGAGAACTTCTTCCCTGAGTCTGCGG
800 810 820 830 840 850 860
TTTCCAATGCTCAGCTGTGAGTTTTTCCCGAGTCTCTCCAAACAGAGAACTTCTTCCCTGAGTCTGCGG
800 810 820 830 840 850 860
```

TTTCCAATGTCAGCTGTGAGTGTGTTTTCCTCCAGTCATCTCCAAACAGGAAGTCTTCTCCCTGAGTGCCTGCGG 800 810 820 830 840 850 860

AGAAGGCTGAGCAAAACCCACAGAGGATCCGACGGGTTTCCACCTCAGAACAAATCGGTTGGCGGTGGG 870 880 890 900 910 920 930

AGAAGGCTGAGCAAAACCCACAGAGGATCCGACGGGTTTCCACCTCAGAACAAATCGGTTGGCGGTGGG 870 880 890 900 910 920 930

940 950 960 970 980 990 1000

GGCGGAAAGAGTGGCGTTGGGATCTGAATCTTCACCATTCACCACCTTTTGGTGAGACCTGGGGTGA 940 950 960 970 980 990 1000

GGCGGAAAGAGTGGCGTTGGGATCTGAATCTTCACCATTCACCACCTTTTGGTGAGACCTGGGGTGA 940 950 960 970 980 990 1000

1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150

GCTCTAGGCTGGGAGCTCCTGAGAGAGCCCTACCTCGGGCCCTTCCACACCTTTGGCAATGTTCTTTT 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150

GCTCTAGGCTGGGAGCTCCTGAGAGAGCCCTACCTCGGGCCCTTCCACACCTTTGGCAATGTTCTTTT 1010 1020 1030 1040 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150

1090 1100 1110 1120 1130 1140 1150

GCCTGGAATAATTAAGTATATGTTTGAACCTTTGAACCTGAACAAATCTCTTTTCGGCTAGGCTTTAT 1090 1100 1110 1120 1130 1140 1150

GCCTGGAATAATTAAGTATATGTTTGAACCTTTGAACCTGAACAAATCTCTTTTCGGCTAGGCTTTAT 1090 1100 1110 1120 1130 1140 1150

1160 1170 1180 1190 1200 1210 1220

GATTTCGAATGCTGTGTAATTAAGAGGCTCTCTACAAGTACTGATAATGAACATGTAAGCAATGCAC 1160 1170 1180 1190 1200 1210 1220

GATTTCGAATGCTGTGTAATTAAGAGGCTCTCTACAAGTACTGATAATGAACATGTAAGCAATGCAC 1160 1170 1180 1190 1200 1210 1220

1230 1240 1250 1260 1270 1280 1290

CACCTTAAGTTACATTCATATGATCTTATTTGATTTTCACTAGGCATAGGAGTAGGAGCTAATAATA 1230 1240 1250 1260 1270 1280 1290

CACCTTAAGTTACATTCATATGATCTTATTTGATTTTCACTAGGCATAGGAGTAGGAGCTAATAATA 1230 1240 1250 1260 1270 1280 1290

1300 1310 1320 1330 1340 1350 1360

CGTTTATTTACTAGAGTTAATCGGAATCAGATATATAACTCTTTTCAGGTACAAAGAACATAATAA 1300 1310 1320 1330 1340 1350 1360

CGTTTATTTACTAGAGTTAATCGGAATCAGATATATAACTCTTTTCAGGTACAAAGAACATAATAA 1300 1310 1320 1330 1340 1350 1360

1370 1380 1390 1400 1410 1420 1430 1440

TCTGGTTTCTGATGTTTATTCAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTATTTGGCTGTAG 1370 1380 1390 1400 1410 1420 1430 1440

TCTGGTTTCTGATGTTTATTCAGTACTACAGCTGCTTCTAATCTTAGTTGACAGTATTTGGCTGTAG 1370 1380 1390 1400 1410 1420 1430 1440

1450 1460 1470 1480 1490 1500 1510

TGTAGCACAGTGTCTCTGGGTACACGGCGGCTCAGCACAGCAGCTTTGAGTTTGGTACTACGTATPCC 1450 1460 1470 1480 1490 1500 1510

TGTAGCACAGTGTCTCTGGGTACACGGCGGCTCAGCACAGCAGCTTTGAGTTTGGTACTACGTATPCC 1450 1460 1470 1480 1490 1500 1510

1520 1530 1540 1550 1560 1570 1580

ACATTTTACACATCACAGAATGAGGCATGCGGCTCTCTCTGGCAATTTATTCAGTACGTATPCC 1520 1530 1540 1550 1560 1570 1580

ACATTTTACACATGACAGAAATGAGGCATGCGGCTCTCTCTGGCAATTTATTCAGTACGTATPCC 1520 1530 1540 1550 1560 1570 1580

1590 1600 1610 1620 1630 1640 1650

GCTTTGGTGGCAGAGCTCATGCTCCACTTCATAGCTATGATTTCTTAAACATCACACTGCATTAGAGTTGA 1590 1600 1610 1620 1630 1640 1650

GCTTTGGTGGCAGAGCTCATGCTCCACTTCATAGCTATGATTTCTTAAACATCACACTGCATTAGAGTTGA 1590 1600 1610 1620 1630 1640 1650

1660 1670 1680 1690 1700 1710 1720

ATAATAAAATTTTCATGTTGAGCAGAAATATTCATGTTTACAGATGTAATGAGTCCCGACCATGTTGCA 1660 1670 1680 1690 1700 1710 1720

ATAATAAAATTTTCATGTTGAGCAGAAATATTCATGTTTACAGATGTAATGAGTCCCGACCATGTTGCA 1660 1670 1680 1690 1700 1710 1720

ATAATAAAATTTTCATGTTGAGCAGAAATATTCATGTTTACAGATGTAATGAGTCCCGACCATGTTGCA 1660 1670 1680 1690 1700 1710 1720

1660 1670 1680 1690 1700 1710 1720

1730 1740 1750 1760 1770 1780 1790 1800

CTGTTCAAGCCCCAAGGAGAGAGCAGGGAACAAGTCTTTACCTTTTATATTTTTCATTTCTAGTGGAGA 1730 1740 1750 1760 1770 1780 1790 1800

CTGTTCAAGCCCCAAGGAGAGAGCAGGGAACAAGTCTTTACCTTTTATATTTTTCATTTCTAGTGGAGA 1730 1740 1750 1760 1770 1780 1790 1800

1810 1820 1830 1840 1850 1860 1870

GATGACAATAAGCAAAATGAGCAAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGA 1810 1820 1830 1840 1850 1860 1870

GATGACAATAAGCAAAATGAGCAAGAAAGATATACAACATCAGGAAATCATGGGTGTTGTGAGAAGCAGAGA 1810 1820 1830 1840 1850 1860 1870

1880 1890 1900 1910 1920 1930 1940

TCAGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGAAATAGATATATTTAGTGGAG 1880 1890 1900 1910 1920 1930 1940

TCAGGCAAGTCACTCTGGGGCTGACACTTGAGCAGAGACATGAAGAAATAGATATATTTAGTGGAG 1880 1890 1900 1910 1920 1930 1940

1950 1960 1970 1980 1990 2000 2010

CAGTATTTCCAGGCAAACTGAGTGGCCCTGGCAAGTTGGATTAAAGCGGGTTTCTCAGACACTACTCAT 1950 1960 1970 1980 1990 2000 2010

CAGTATTTCCAGGCAAACTGAGTGGCCCTGGCAAGTTGGATTAAAGCGGGTTTCTCAGACACTACTCAT 1950 1960 1970 1980 1990 2000 2010

2020 2030 2040 2050 2060 2070 2080

GTGTGTGTGTGG 2020 2030 2040 2050 2060 2070 2080

GTGTGTGTGTGG 2020 2030 2040 2050 2060 2070 2080

2090 2100 2110 2120 2130 2140 2150 2160

GCAGTATCCTGCTCCCTCCCTACCTAGTGTGTAGGACACTCCCGGCTTGAACACCAAAATGCTCT 2090 2100 2110 2120 2130 2140 2150 2160

GCAGTATCCTGCTCCCTCCCTACCTAGTGTGTAGGACACTCCCGGCTTGAACACCAAAATGCTCT 2090 2100 2110 2120 2130 2140 2150 2160

2170 2180 2190 2200 2210 2220 2230

AACTTTGCGCACATGTCACCTAGTAGACAACTCCCTGTTAAAGAGCTCGGGTTGAAAGAAATTAACAAGTA 2170 2180 2190 2200 2210 2220 2230

AACTTTGCGCACATGTCACCTAGTAGACAACTCCCTGTTAAAGAGCTCGGGTTGAAAGAAATTAACAAGTA 2170 2180 2190 2200 2210 2220 2230

2240 2250 2260 2270 2280 2290 2300

GTGCTGGGAGTAGAGGCCAAGATAGTAAATGGGCTCAGAAGAGGAGCCCAACAAGGTTGTGCAAGGG 2240 2250 2260 2270 2280 2290 2300

GTGCTGGGAGTAGAGGCCAAGATAGTAAATGGGCTCAGAAGAGGAGCCCAACAAGGTTGTGCAAGGG 2240 2250 2260 2270 2280 2290 2300

2310 2320 2330 2340 2350 2360 2370

CCTGTAGGCTGTGGTGTGAATTTAGCCAAAGAGTAACAGTATCTGTACAGGCTTTTAAAGAGTTGCTCT 2310 2320 2330 2340 2350 2360 2370

CCTGTAGGCTGTGGTGTGAATTTAGCCAAAGAGTAACAGTATCTGTACAGGCTTTTAAAGAGTTGCTCT 2310 2320 2330 2340 2350 2360 2370

2380 2390 2400 2410 2420 2430 2440

GGCTGCTATGTGGAAGCAGATGAGGAGGAGCAACAGTAAGAGGAGGAGCCCGCAGGAGGTTGTACAC 2380 2390 2400 2410 2420 2430 2440

GGCTGCTATGTGGAAGCAGATGAGGAGGAGCAACAGTAAGAGGAGGAGCCCGCAGGAGGTTGTACAC 2380 2390 2400 2410 2420 2430 2440

2450 2460 2470 2480 2490 2500 2510 2520

AGTCCAGGCAAGAGTAGTGGGTGGGTGGGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2450 2460 2470 2480 2490 2500 2510 2520

AGTCCAGGCAAGAGTAGTGGGTGGGTGGGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2450 2460 2470 2480 2490 2500 2510 2520

2530 2540 2550 2560 2570 2580 2590

TATATCTGAAGGAAGTGTGCTGAAGATCTATGTTGTGAGAGAAAGAGAAATTTGGTGGGTGTAGTA 2530 2540 2550 2560 2570 2580 2590

TATATCTGAAGGAAGTGTGCTGAAGATCTATGTTGTGAGAGAAAGAGAAATTTGGTGGGTGTAGTA 2530 2540 2550 2560 2570 2580 2590

2600 2610 2620 2630 2640 2650 2660
GCTCATGCCAAGGAGGCCAAGGAGAGCATCTCTGAGCTCAGGAGTTCAAGACCAGCGCTCGGCAACAC
|||||
GCTCATGCCAAGGAGGCCAAGGAGAGCATCTCTGAGCTCAGGAGTTCAAGACCAGCGCTCGGCAACAC
2600 2610 2620 2630 2640 2650 2660
2670 2680 2690 2700 2710 2720 2730
ACCAAAACCCCTTCTCTACAAAAATACAAAAATAGCTGGTGTGGTGCGATGCACCTGTGATCCTAGCTA
|||||
ACCAAAACCCCTTCTCTACAAAAATACAAAAATAGCTGGTGTGGTGCGATGCACCTGTGATCCTAGCTA
2670 2680 2690 2700 2710 2720 2730
2740 2750 2760 2770 2780 2790 2800
CTCGGAGGCTCAGCTGGAGGCTATTGCTTGAGCCAGGAAGTTGAGGCTCAGTGAGCCATGACTGTGCCA
|||||
CTCGGAGGCTCAGCTGGAGGCTATTGCTTGAGCCAGGAAGTTGAGGCTCAGTGAGCCATGACTGTGCCA
2740 2750 2760 2770 2780 2790 2800
2810 2820 2830 2840 2850 2860 2870 2880
CTGTACTTCAGCCTAGGTGACAGACAAGACCTGTCTCCCTGACCCCTTGAAAAAGAGAAGATTAAAGT
|||||
CTGTACTTCAGCCTAGGTGACAGACAAGACCTGTCTCCCTGACCCCTTGAAAAAGAGAAGATTAAAGT
2810 2820 2830 2840 2850 2860 2870 2880
2890 2900 2910 2920 2930 2940 2950
TGACTTTGTTCTTTAATTTAATTTAATTTGCGCTGAGCAGTGGGTAAATTGGCAATGCCATTTCTGAGATGCT
|||||
TGACTTTGTTCTTTAATTTAATTTAATTTGCGCTGAGCAGTGGGTAAATTGGCAATGCCATTTCTGAGATGCT
2890 2900 2910 2920 2930 2940 2950
2960 2970 2980 2990 3000 3010 3020
GAAGCAGAGGAAGAGCAGTTTGGGTAATCAAGGATCTGCATTTGGGACATCTTAAGTTTGAGATTCCA
|||||
GAAGCAGAGGAAGAGCAGTTTGGGTAATCAAGGATCTGCATTTGGGACATCTTAAGTTTGAGATTCCA
2960 2970 2980 2990 3000 3010 3020
3030 3040 3050 3060 3070 3080 3090
GTCAAGCTTCCAAAGTGGTGAGGCCACATAGGCAGTTTCAGTGTAAAGAAATTCAGGACCAAGGCTGGCGCGGTG
|||||
GTCAAGCTTCCAAAGTGGTGAGGCCACATAGGCAGTTTCAGTGTAAAGAAATTCAGGACCAAGGCTGGCGCGGTG
3030 3040 3050 3060 3070 3080 3090
3100 3110 3120 3130 3140 3150 3160
GCTCACTTCTGTAATCCCGACACTTTGGTGGCTGAGGCGAGGTAGATCATTTGAGTTCAGGAGTTTGAGACAA
|||||
GCTCACTTCTGTAATCCCGACACTTTGGTGGCTGAGGCGAGGTAGATCATTTGAGTTCAGGAGTTTGAGACAA
3100 3110 3120 3130 3140 3150 3160
3170 3180 3190 3200 3210 3220 3230 3240
GCTTGGCCAAACATGGTGAACCCCATGTCTACTTAAATAATACAAAAATAGCCCTGGTGGTGGCGCAGCCCT
|||||
GCTTGGCCAAACATGGTGAACCCCATGTCTACTTAAATAATACAAAAATAGCCCTGGTGGTGGCGCAGCCCT
3170 3180 3190 3200 3210 3220 3230 3240
3250 3260 3270 3280 3290 3300 3310
ATAGTCCAGGTTTTCCAGGAGGTTAGTGGAGAGATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAAGCT
|||||
ATAGTCCAGGTTTTCCAGGAGGTTAGTGGAGAGATCCCTTGAACCCAGGAGGTGCAGGTTGCAGTGAAGCT
3250 3260 3270 3280 3290 3300 3310
3320 3330 3340 3350 3360 3370 3380
GAGATTGTGCCACTGCACCTCCAGCCCTGGGTGATAGAGTGAAGTCTGTCTCAAAAAAATTTTTTTTTT
|||||
GAGATTGTGCCACTGCACCTCCAGCCCTGGGTGATAGAGTGAAGTCTGTCTCAAAAAAATTTTTTTTTT
3320 3330 3340 3350 3360 3370 3380
3390 3400 3410 3420 3430 3440 3450
AAAAAACTGAAGGAATTTATCCTCAGGATTTGGGTCTAATTTGGCCCTGAGCACCACCTCCTGAGTTCAAC
|||||
AAAAAACTGAAGGAATTTATCCTCAGGATTTGGGTCTAATTTGGCCCTGAGCACCACCTCCTGAGTTCAAC
3390 3400 3410 3420 3430 3440 3450

3460 3470 3480 3490 3500 3510 3520
TACCATGGCTAGACACACCTTTAATTTCTAGAAATCCACAGCTTTAGTGAGTCTGTCTAATCATGAGTA
|||||
TACCATGGCTAGACACACCTTTAATTTCTAGAAATCCACAGCTTTAGTGAGTCTGTCTAATCATGAGTA
3460 3470 3480 3490 3500 3510 3520
3530 3540 3550 3560 3570 3580 3590 3600
TTGGAATFAGGATCTGGGGCAGTGAGGGGTGGCAGCGTGGCAGAGAAAACACACAGGAAGAGC
|||||
TTGGAATFAGGATCTGGGGCAGTGAGGGGTGGCAGCGTGGCAGAGAAAACACACAGGAAGAGC
3530 3540 3550 3560 3570 3580 3590 3600
3610 3620 3630 3640 3650 3660 3670
ACCCAGGACTGTCTATGGAAGAAAGACAGGACTGCAACTCACACCTTCAAAAATGAGGACACACAGCT
|||||
ACCCAGGACTGTCTATGGAAGAAAGACAGGACTGCAACTCACACCTTCAAAAATGAGGACACACAGCT
3610 3620 3630 3640 3650 3660 3670
3680 3690 3700 3710 3720 3730 3740
GATGGTATGAGTTGATGACAGGTGTGGAGCCTCAACATCCTCTCCCTTCTACACATGATGGTTAAGGCC
|||||
GATGGTATGAGTTGATGACAGGTGTGGAGCCTCAACATCCTCTCCCTTCTACACATGATGGTTAAGGCC
3680 3690 3700 3710 3720 3730 3740
3750 3760 3770 3780 3790 3800 3810
TGTTGCTCTGTCTCCAGGTTCAACATCTCTGACCTACCTTTCATGGGTGCCCTCAGAGCAGGACCTTGGTCT
|||||
TGTTGCTCTGTCTCCAGGTTCAACATCTCTGACCTACCTTTCATGGGTGCCCTCAGAGCAGGACCTTGGTCT
3750 3760 3770 3780 3790 3800 3810
3820 3830 3840 3850 3860 3870 3880
TTCCTTCTTTGAAGCTTTGGGCTACGTGGATGACCACTGTCTGTTCTATGATNNTGAGAGTCCCGTGT
|||||
TTCCTTCTTTGAAGCTTTGGGCTACGTGGATGACCACTGTCTGTTCTATGATNNTGAGAGTCCCGTGT
3820 3830 3840 3850 3860 3870 3880
3890 3900 3910 3920 3930 3940 3950 3960
GGAGCCCGAACTCCATGGGTTCCAGTAGAATTTCAAGCCAGATGTGGTGCAGCTGAGTCAAGTCTGAA
|||||
GGAGCCCGAACTCCATGGGTTCCAGTAGAATTTCAAGCCAGATGTGGTGCAGCTGAGTCAAGTCTGAA
3890 3900 3910 3920 3930 3940 3950 3960
3970 3980 3990 4000 4010 4020 4030
AGGTGGATCACATGTTCACTGTGACTTCTGACATTTATGAAAAATCACACCAAGGATGATGTG
|||||
AGGTGGATCACATGTTCACTGTGACTTCTGACATTTATGAAAAATCACACCAAGGATGATGTG
3970 3980 3990 4000 4010 4020 4030
4040 4050 4060 4070 4080 4090 4100
GAGAGGGGCTCACCTTCTGAGGTTGTGAGCTTTTCATCTTTTCATGCACTTTGAGGAAACAGCTGG
|||||
GAGAGGGGCTCACCTTCTGAGGTTGTGAGCTTTTCATCTTTTCATGCACTTTGAGGAAACAGCTGG
4040 4050 4060 4070 4080 4090 4100
4110 4120 4130 4140 4150 4160 4170
AAGTCTGAGTCTTGTGGGACGAGGAGGAAGTAATTTGCTTCTGAGATCATTTGGTCTCTGGGAT
|||||
AAGTCTGAGTCTTGTGGGACGAGGAGGAAGTAATTTGCTTCTGAGATCATTTGGTCTCTGGGAT
4110 4120 4130 4140 4150 4160 4170
4180 4190 4200 4210 4220 4230 4240
GGTGAATAGGGACCTTTCCTTTGGTGGTGAAGTAATTAACAGGCTGGGATTTTCCAGAGTCCCACACCTCG
|||||
GGTGAATAGGGACCTTTCCTTTGGTGGTGAAGTAATTAACAGGCTGGGATTTTCCAGAGTCCCACACCTCG
4180 4190 4200 4210 4220 4230 4240
4250 4260 4270 4280 4290 4300 4310 4320
CAGTCTACCTTGGGCTGTAATCAAGAAACACAGTACCGAGGCTACTGGAAGTACGGGTATATGGG
|||||
CAGTCTACCTTGGGCTGTAATCAAGAAACACAGTACCGAGGCTACTGGAAGTACGGGTATATGGG
4250 4260 4270 4280 4290 4300 4310 4320
4330 4340 4350 4360 4370 4380 4390
4400 4410 4420 4430 4440 4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560 4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680 4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800 4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920 4930 4940 4950 4960 4970 4980 4990 5000

CGTCTGGCACCCCTAGTCTATTGGAGTCAATCAGTGGAAATTGCTGTTTTTTTGTCTGTCATCTTTGTTCTATTGGAAATTT
6050 6060 6070 6080 6090 6100 6110 6120
TGTTTCATATATTAAGGAAGAGCGAGGGTCAAGTGAGTAGGAACAAGGGGAAGTCTCTTTAGTACCTCTGC
6130 6140 6150 6160 6170 6180 6190
TGTTTCATATATTAAGGAAGAGCGAGGGTCAAGTGAGTAGGAACAAGGGGAAGTCTCTTTAGTACCTCTGC
6130 6140 6150 6160 6170 6180 6190
6200 6210 6220 6230 6240 6250 6260
CCGAGGACACAGTGGAGAGGGGAGAGGGGATCTGGCATCCATGGGAAGCATTTTCTCATTTATATCT
6270 6280 6290 6300 6310 6320 6330
TTGGGGACACCAAGAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCAAGATAAGAGTCTCTAAATTCAC
6340 6350 6360 6370 6380 6390 6400
TTGGGGACACCAAGAGCTCCCTGGGAGACAGAAAATAATGGTTCTCCCAAGATAAGAGTCTCTAAATTCAC
6340 6350 6360 6370 6380 6390 6400
6410 6420 6430 6440 6450 6460 6470 6480
CACTGTGGCTATTCTCAGAACCCAACTATTTCGACAGAGCTGTTAAGGTAGTACAGGGCTTTGAGGTTGAGAAGT
6490 6500 6510 6520 6530 6540 6550
ACCCATGAGGTCTTAAACAGCGAGAGCAAGCAAAATGCTTTAGGGTGTCAAAGGAAGATGATCATCATTCAGC
6560 6570 6580 6590 6600 6610 6620
TGCGGATCAAGATAGGCTTCTGATCTTCAAGGAAGAGCTGGATCCCATAGCTAGCTTGAAGATGATGGG
6630 6640 6650 6660 6670 6680 6690
AGTCTACACAGCGAGCAACCATGCCAAGTAGGAGATATAAGGCATACCTGGAGATTAGAAATATTAC
6700 6710 6720 6730 6740 6750 6760
TGTTACCTTAACCTGAGTTTGGCTAGCTATCACTACCAATATTGCAATTTCTACCCCTGGAACATCTGTGT
6770 6780 6790 6800 6810 6820 6830 6840
GTAGGGAAGAAGCAATCAGAAAGCCAGCTCATACAGAGTCCAAAGGCTCTTTGGGATATTGGGTATGA
6850 6860 6870 6880 6890 6900 6910
TCAGTGGGTGTCTATTGAAGGATCCCTAAGAAAGGAGGACACGATCTCCCTTATATGSGAATGTTGTGA
6920 6930 6940 6950 6960 6970 6980
TCAGTGGGTGTCTATTGAAGGATCCCTAAGAAAGGAGGACACGATCTCCCTTATATGSGAATGTTGTGA
6990 7000 7010 7020 7030 7040 7050
AGAGTTAGATGAGAGGTGAGGACACCAAGTGTAGAAAGCAATTAAGCATTTCCAGATGAGAGATAATGGTTCT

6990 7000 7010 7020 7030 7040 7050
TGAATCCAATAGTGCCTCCAGGTCTAAATTCAGATGSGTGAATGAGAAATAGGAGAGAGGCAAG
6990 7000 7010 7020 7030 7040 7050
TGAATCCAATAGTGCCTCCAGGTCTAAATTCAGATGSGTGAATGAGAAATAGGAGAGAGGCAAG
6990 7000 7010 7020 7030 7040 7050
7060 7070 7080 7090 7100 7110 7120
ATGTGCTAGTGTGATGCTCTTTCTGCTGCTCTCTCCACAGGAGGACCATGGGCACTACGT
7060 7070 7080 7090 7100 7110 7120
ATGTGCTAGTGTGATGCTCTTTCTGCTGCTCTCTCCACAGGAGGACCATGGGCACTACGT
7060 7070 7080 7090 7100 7110 7120
7130 7140 7150 7160 7170 7180 7190 7200
CTTAGCTGAACGTGAGTGACACGCGCTGCAGACTCAGCTGCTGGGAGGAGACAAAACCTAGACACTCAAGA
7130 7140 7150 7160 7170 7180 7190 7200
CTTAGCTGAACGTGAGTGACACGCGCTGCAGACTCAGCTGCTGGGAGGAGACAAAACCTAGACACTCAAGA
7130 7140 7150 7160 7170 7180 7190 7200
7210 7220 7230 7240 7250 7260 7270
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTTAAACATAGAAATTCCTGAGACT
7210 7220 7230 7240 7250 7260 7270
GGGAGTGCATTTATGAGCTCTTCATGTTTCAGGAGAGAGTTGAACCTTAAACATAGAAATTCCTGAGACT
7210 7220 7230 7240 7250 7260 7270
7280 7290 7300 7310 7320 7330 7340
CCTTGATTTTGAAGCTCTCTGTTTCATTTCTCAAAAAGATTTCCCAATTTAGTTTCTGAGTTCTCTGCATGC
7280 7290 7300 7310 7320 7330 7340
CCTTGATTTTGAAGCTCTCTGTTTCATTTCTCAAAAAGATTTCCCAATTTAGTTTCTGAGTTCTCTGCATGC
7280 7290 7300 7310 7320 7330 7340
7350 7360 7370 7380 7390 7400 7410
CGGTGATCCCTTAGCTGAGCTCTCCCTGGAAGTCTCTCATGAACCTCAAGCTGCATCTAGAGCTTCC
7350 7360 7370 7380 7390 7400 7410
CGGTGATCCCTTAGCTGAGCTCTCCCTGGAAGTCTCTCATGAACCTCAAGCTGCATCTAGAGCTTCC
7350 7360 7370 7380 7390 7400 7410
7420 7430 7440 7450 7460 7470 7480
TTTCATTTCTCCCTGACCTCAGACACATACACCTATGCTCATTTCTTCTTCTTGGAGAGGACTCCTT
7420 7430 7440 7450 7460 7470 7480
TTTCATTTCTCCCTGACCTCAGACACATACACCTATGCTCATTTCTTCTTCTTGGAGAGGACTCCTT
7420 7430 7440 7450 7460 7470 7480
7490 7500 7510 7520 7530 7540 7550 7560
AAATTTGGGGACTTACATGATTTCAATTTAAACATCTGAGAAAGCTTTGAACCTCGGAGCTGGCTAGTCT
7490 7500 7510 7520 7530 7540 7550 7560
AAATTTGGGGACTTACATGATTTCAATTTAAACATCTGAGAAAGCTTTGAACCTCGGAGCTGGCTAGTCT
7490 7500 7510 7520 7530 7540 7550 7560
7570 7580 7590 7600 7610 7620 7630
AACCTTACAGATTTTACACATGATTTCAATTTCTGGACCTTCAACCTTTCTTCTTGAATCTCTC
7570 7580 7590 7600 7610 7620 7630
AACCTTACAGATTTTACACATGATTTCAATTTCTGGACCTTCAACCTTTCTTCTTGAATCTCTC
7570 7580 7590 7600 7610 7620 7630
7640 7650 7660 7670 7680 7690 7700
TCTGTGTACCCAGTAACTCATCTGTCAACAGCTTTGGGGATCTTCCATCTCATTTGATGTGAGTCTGA
7640 7650 7660 7670 7680 7690 7700
TCTGTGTACCCAGTAACTCATCTGTCAACAGCTTTGGGGATCTTCCATCTCATTTGATGTGAGTCTGA
7640 7650 7660 7670 7680 7690 7700
7710 7720 7730 7740 7750 7760 7770
CAGCTATCAAGGCTGTACACTGCAGGAATGGAAGAGGACCTGTCCAGAAAGACATCATGCTATCTGTG
7710 7720 7730 7740 7750 7760 7770
CAGCTATCAAGGCTGTACACTGCAGGAATGGAAGAGGACCTGTCCAGAAAGACATCATGCTATCTGTG
7710 7720 7730 7740 7750 7760 7770
7780 7790 7800 7810 7820 7830 7840
GGTAGTATGATGGGTGTTTATAGCAGGTAGGAGCAAAATATCTTTGAAAGGGTGTGTGAAGAGGTGTTTTC
7780 7790 7800 7810 7820 7830 7840
GGTAGTATGATGGGTGTTTATAGCAGGTAGGAGCAAAATATCTTTGAAAGGGTGTGTGAAGAGGTGTTTTC


```
270 280 290 300 310 320
480
GCGGGGTG-GAAAAATC-GAAATAGCTTTTCTTTCGCGCTGGGAG-TTTC--CTAACTTTGGAGACCT
|||||
TCATGGTGCCTCAGAGCAGGACCTTGGTCTTCCCTTG---TTTGAAGCTTTGGGCTA--CGTGGATGACCA
330 340 350 360 370 380
540 550 560 570 580 590 600
GCTCAAC-----CCTATCCGCAAGCCCTCTCCCTACTTCTCGGTCCAGACCCCGTGGAGGTGCTCTACCA
|||||
GCTGTTCGTGTTCTATGATCATGATGAGTCCGCGT--GTGGAGC-CCCGAACTCAT--GGGTTT-----CCA
390 400 410 420 430 440 450
610 620 630 640 650 660 670
CT-GAATGTC-AGATAGGGTCCCTCCGCCAGACCTGCCCTCCCGCGCTGCCCGGTCTGCGGAG-
|||||
GTAGAAATTTCAAGCCAGATGGCT-CAAGCTGAGTCAG-AGTCTGAAAGGGTGGGATCACATGTTCACTGT
460 470 480 490 500 510 520
680 690 700 710 720 730 740
TGACTTTTGGGA----ACGCCCACTCCCTCCCACTAGATGCTTTTAAATAATCTCTAGTCTCTCA
|||||
TGACTTCGGACTATTATGGAATAC-ACCACAGCAGGA--GTCCCAACCCCTGCAGGTCACTCTGGG
530 540 550 560 570 580
750 760 770 780 790 800 810
CT-TGAGCTGAGCTAAGCTCGGGCTCCTTGAAC--CTGGAAT-CGGGT-TTATTTCCAATGTGAGCTGTG
|||||
CTGTGAATGCAAGAGACACACATACCGAGGGCTACTGGAAGTACGGGTATGATGGGAGGACACACCT-TG
590 600 610 620 630 640 650 660
820 830 840 850 860 870
CAGTTTTCCTCCAGTATCTCCAAACAGGA----AGTTCTTCCCTGAGTGTCTGCGGAGAGGT-GAG--
|||||
AA---TTCTGCCCTGACA-CACTGGATTGGAGAGCAGCAGAACCCAG-GGCCCTGGCCACCRAGCTGGAGT
670 680 690 700 710 720
880 890 900 910 920 930
-CAAACCCACAGCA--GGATC-CCGA---CGGGTTTCCACCT-CAGAACGAATCGTTGGCGGTGGGG
|||||
GGAAGGCAAGATTCGGGCCAGGCAAGAGAGG--CCTACCTGGAGAGGACTCCCTGCACAGCTGCAG
730 740 750 760 770 780 790
940 950 960 970 980 990 1000 1010
CGCGAAAGAGTGGGATCTGAATCTTCCACCATCCACCCTTGGTGAGACCTGGGGTGAGG
|||||
C---AGTTGCTGGAGCTGGGA---GAGGTGT-----TTGGACCAAC---AAGTCCCTCTTGTGTGAAG
800 810 820 830 840 850
1020 1030 1040 1050 1060 1070
TCTC-TAGGGTGGGAGCTCCTGAGAGA--GGCTAG---CTCGGGCTTTCCCACT---CTTGGCAATTG
|||||
TGACATCATGTGA-CCTCTTCACTGACCACCTCTAGGCTGTCGGGCTTGAACCTACTACCCCCAGACATC
860 870 880 890 900 910 920
1080 1090 1100 1110 1120 1130 1140
TTCTTT---TTCCCTGGAATTAAGTATATGTTAGTTTGA--CGTTTGAACCTGAACAAATCTCTTTGG
|||||
ACCATGAAGTGGCT-GAAGGATAGCAGCAATGGATGCCAAGGAGTTTCAACCTAA-AGACGTATTGCCA
930 940 950 960 970 980 990
1150 1160 1170 1180 1190 1200
CTAGCTTTATTTGTAATGCTGTGTAA--TTAAGAGCCCTCTCT-ACAAAGTACTGATATGATTAATGAACA
|||||
ATGGGAT--GGGACTACCA-GGGCTGGATAACCTTGGCTGTACCCCTCGGGGAGAGCAGA-GATATACG
1000 1010 1020 1030 1040 1050 1060
1210 1220 1230 1240 1250 1260 1270
TG-TAAGCAATGCACTCTCTAAGTTA--CATTCATATCTGATCTTATTTGATTTTCACT-AGGCA---T
|||||
TGCCAGTGGAGACACCCAGCCCTGGATGACCCCTCAT-TGTGATCTGGGAGCCCTCACCGTCTGGCACCT
1070 1080 1090 1100 1110 1120 1130
```

```
1280 1290 1300 1310 1320 1330 1340
AGGAGGTAGGAGCTAATAATACGTTTATTTTACTAGAAGTTAACTGGAAATCAGATATATAACTTTTC
|||||
AGTCA--TTGGAGTCATCAGTGAATTTGTTTGTGCTCATCTTG--TTC--ATT--GGAATTTGTTTC
1140 1150 1160 1170 1180 1190
1350 1360 1370 1380 1390 1400 1410
AGGTTACAAGAAGACATAAATACTGGTTTCTTCTGATGTTTAACTCAAGTACTACAGCT--GCTTCTAATCTTA
|||||
ATAATAATAAG-----GAAGAGCAGGCTTCAAGAGGAGCCACTGGGCACTACGTTTACGT--GAACGTGA
1200 1210 1220 1230 1240 1250 1260
1420 1430 1440 1450 1460 1470
GTTCAAGATGATTTG-----CCCTGAGTGTAGCACAGTGTCT-TGTGGTACACGCCGCCCT-CA-GC
|||||
G-TGACACGAGCGCTGCGAGACTCACTGT--GGGAAGGAGCAAAACTTAGAGACTCA--AAGAGGAGTGCATTT
1270 1280 1290 1300 1310 1320 1330
1480 1490 1500 1510 1520 1530 1540
A-CAGCACTTGTGCTTTGTGCTACTAGCTGTATCCACATTTTACACATGACAAGATGAGGCATG--GCACGG
|||||
ATGAGCTTCTCATGTTCAGGAGAGAGTTGAACC-----TAAACAT-AGA--AAT--TGCCTGACGAATC
1340 1350 1360 1370 1380 1390
1550 1560 1570 1580 1590 x 1610 1620
CCTGCTTCCCTGGCAAAATTTATCAATGTACACTGGCTTGTGGCAGAGCTCATGTCTCCACTTCATAGC
|||||
CTTGATTTAGCCTTCTCTGTTTATTTCTTCAAAAAGATTT--CCCCA
1400 1410 1420 1430 1440
TATGATTTTAAACATCACACTGATAGAGTTGAATAATAAATTTTCATGTTGAGCAGAAATATTTCATTCG
1630 1640 1650 1660 1670 1680 1690
TTTA
```

6. US-09-497-957-1 (1-10825)

US-09-497-957-1 Sequence 11, Application US/09497957

Initial Score = 676 Optimized Score = 785 Significance = -0.00
Residue Identity = 58% Matches = 894 Mismatches = 471
Gaps = 167 Conservative Substitutions = 0

```
40 50 60 70 80 90 100 110
TTTGAAATCATAAATATTTAAATATCTAAAGTTCAGATCAGAACATTGCGAAGCTACTTTCCCAATCAAC
120 130 140 150 160 170 180
AACACCCCTTTCAGGATTTAAACCAACCAAGGGGACACTGGATCACCTAGTGTTCACAGCAGGTACCTTCTG
|||||
GGGACACTGGATCACTACCTAGTGTTCACAGCAGGTACCTTCTG
x 10 20 30 40
190 200 210 220 230 240 250
CTGTAGGAGAGAGAACTAAAGTCTTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCC
|||||
CTGTAGGAGAGAGAACTAAAGTCTTGAAGACCTGTGCTTTTACCAGGAAGTTTACTGGGCATCTCC
50 60 70 80 90 100 110
260 270 280 290 300 310 320
TGAGCTTAGCAATAGCTGTAGGTGACTTCTGGAGCCATCCCGTTCCTCCGCCCCCAAGAGCGGAG
120 130 140 150 160 170 180
TGAGCTTAGCAATAGCTGTAGGTGACTTCTGGAGCCATCCCGTTCCTCCGCCCCCAAGAGCGGAG
|||||
ATTTAAGGGAGCTGGGCCAGAGCTGGGGAATGGGCCCGGAGCCGCGCTTCTCTCTGATG
330 340 350 360 370 380 390
ATTTAAGGGAGCTGGGCCAGAGCTGGGGAATGGGCCCGGAGCCGCGCTTCTCTCTGATG
|||||
ATTTAAGGGAGCTGGGCCAGAGCTGGGGAATGGGCCCGGAGCCGCGCTTCTCTCTGATG
190 200 210 220 230 240 250 260
400 410 420 430 440 450 460 470
```

```
CTTTTCAGACCGGGCTTCCTGAGGGGCGCTTCTGCTGAGTCCGAGGGCTCGGGGGAACCTAGGGGCGCG
|||||
CTTTTCAGACCGGGCTTCCTGAGGGGCGCTTCTGCTGAGTCCGAGGGGCGCTTCTGCTGAGTCCGAGGGGCGCG
270 280 290 300 310 320
480 490 500 510 520 530
CGGGGGTG-GAAAACTAGCTTTCTTTTGGCTTGGGAG-TTTG--CTAACCTTTGGAGGAGCT
|||||
TCATGGGTGCTCAGACGAGGAGCTTTGGCTTCTTCCCTG---TTTGAAGCTTTGGGCTA--CGTGGATGACCA
330 340 350 360 370 380
540 550 560 570 580 590 600
GCTCAAC-----CCTATCCGGAAGCCCTCTCCCTACTTTCTCGCTCCAGACCCCTGAGGGAGTGCCTACCA
|||||
GCTGTGCTGTCTATGATGATGAGTGCCTG---GTGGAGC-CCCGAACTCCAT--GGGTTT-----CCA
390 400 410 420 430 440 450
610 620 630 640 650 660 670
CT-GAAGTGC-AGATAGGGTCTCTCGCCCGAGGACCTGCCCTCCCGGCTCTCCGGCTCTCGGGAG-
|||||
GTAGAAATTCAGCCAGATGTGGCT-GCAGCTGAGTCAG-AGTCTGAAGGGTGGGATCAGATGTTCACTGT
460 470 480 490 500 510 520
680 690 700 710 720 730 740
TGACTTTTGA-----ACGGCCACTCCCTTCCCGCACTAGAACTGTTTAAATAAATCTGTAATTCCTCA
|||||
TGACTTCTGGACTATTATGAAAAATCAC-AACACAGCAAGGA--GTCCACACCCCTGCAGTCACTCCTGGG
530 540 550 560 570 580
750 760 770 780 790 800 810
CT-TGAGCTGAGTGAAGCTGGGCTCCTTGAAAC---CTGGAACCT-CGGGT-TTATTTCCAATGTCAGCTGTG
|||||
CTGTGAATTCGAAGAAGAACACACTACCGAGGCTACTGGAAGTACGGGTATGATGGCAGGACCACT-TG
590 600 610 620 630 640 650
820 830 840 850 860 870
CAGTTTTTCCCACTATCTCCAAACAGGA----AGTTCTTCCCTGAGTCTGCTCCGAGAGGCT-GAG--
|||||
AA----TTCTGCCCTGACA-CATCTGGATTGGAGAGCAGCAGAACCCAG-GGCTGGCCCAACCAAGCTGGAGT
670 680 690 700 710 720
880 890 900 910 920 930
--CAACCCACAGCA---GGATC-CGCA---CGGGTTTCCACT-CAGAACCAATGCGTTGGGGGGGGGG
|||||
GGAAGGCACAAAGATTCCGGCCAGGACAGACAGG--CTACTCTGGAGGGAGTGCCTCAGCAGCTGCAG
730 740 750 760 770 780 790
940 950 960 970 980 990 1000
CGCGAAAGAGTGGGCTGGAATCTTCAACATTCACCCACTTTTGTGAGACCTGGGGGTGGAGG
|||||
C---AGTTGCTGGAGCTGGGA---GAGGTGTT-----TTGACCAAC---AAGTGCCTCCTTTTGGTGAAG
800 810 820 830 840 850
1020 1030 1040 1050 1060 1070
TCTC-TAGGGTGGAGGCTCCTGAGAGA--GGCCTAC---CTCGGGCCTTCCCGCACT---CTTGGCAATG
|||||
TGACACATCATGTGA-CCTCTTCAGTGACCACTCTACGGTGTGGGGCTTGAACCTACTACCCCAACATC
860 870 880 890 900 910 920
1080 1090 1100 1110 1120 1130 1140
TCTTPT---TTCCCTGGAATAATAGTATAGTTAGTTTGA--CGTTGAGACTGAACAATCTCTCTTTCGG
|||||
ACCATGAAGTGGCT-GAAGGATAGCAGGCAATGGAATGCAAGGAGTTCGAACCTTAA-AGACGTATTGCCCA
930 940 950 960 970 980 990
1150 1160 1170 1180 1190 1200
CTAGGCTTTATGATTTGCAATGCTGTGTGTAA--TTAAGAGCCCTCTCT-ACAAGTACTGATGAATGAACA
|||||
ATGGGGAT--GGGACCTACCA-GGGCTGGATAACCTTGGCTGTACCCCTGGGGAAGAGCAGA-GATATAG
1000 1010 1020 1030 1040 1050 1060
1210 1220 1230 1240 1250 1260 1270
TG-TAAGCAATGCACCTCACCTTCTTAAGTTA--CATTCATATCTGATCTTATTTGATTTTCACT-AGGCA--T
```

```
|||||
TGCCAGGTGGAGACCCAGGCTTGATCAGCCCTCAT-TGTGATCTGGGAGCCCTCACCCTCTGGCACCCCT
1070 1080 1090 1100 1110 1120
1280 1290 1300 1310 1320 1330 1340
AGGGAGGTAGAGCTAATAACGTTTATTTTACTAGAAGTTAACTGGAATTCAGATTAATAACTCTTTTC
|||||
AGTCA--TTGGAGTCATCAGTGAATTCGTGTTTCTGCTCATCTTG--TTC--ATT--GGATTTTGTTC
1140 1150 1160 1170 1180 1190
1350 1360 1370 1380 1390 1400 1410
AGGTTACAAAGCAATAAATATCTGCTTCTGATGTTATTTCAAGTACTACAGCT--GCTTCTTAATCTTA
|||||
ATAATATTAAAG-----GAAGAGGAGGGTTCAAGAGGACCATGGGCACTAGCTTAGCT--GAACGTGA
1200 1210 1220 1230 1240 1250 1260
1420 1430 1440 1450 1460 1470
GTTGACAGTGAATTTG-----CCTGTAGTGTAGCACAGTGTCT-TGTGGTTCACACCGCGGCT-CA-GC
|||||
G-TGACACGACGCTGCAGACTCACTGT-GGGAAGGAGACAAACTAGAGACTCA-AAGAGGGAGTGCATTT
1270 1280 1290 1300 1310 1320 1330
1480 1490 1500 1510 1520 1530 1540
A-CAGCAGCTTTGAGTTTGTGTACTACGTGTATCCACATTTTACACATGACAAAGATGAGGATG--GCACGG
|||||
ATGAGCTCTTCATGTTTTCAGGAGAGAGTTGAACC-----TAAACAT-AGA--AAT--TGCTGACGAACTC
1340 1350 1360 1370 1380 1390
1550 1560 1570 1580 1590 1600 1610 1620
CCTGTTCTTGGCAAAATTAATCAATGCTGACACTGGCTTGTGGCAGAGCTCATGTCTCCACTTCATAGC
|||||
CTTGATTTTACCTTCTGCTGTTCAITTCCTCAAAAGATTT--CCCCA
1400 1410 1420 1430 1440
1630 1640 1650 1660 1670 1680 1690
TATGATTTCTTAACATCACACTGATAGAGTTGAGTTGAATAAATAAATTTTCATCTTGACGAGAAATATTCATTTG
TTTA
US-09-497-957-1 (1-10825)
US-09-497-957-10 Sequence 10, Application US/09497957
Initial Score = 676 Optimized Score = 786 Significance = -0.00
Residue Identity = 58% Matches = 899 Mismatches = 463
Gaps = 173 Conservative Substitutions = 0
40 50 60 70 80 90 100 110
TTTGAATAATCATAAATAATTTAAATATCTAAAGTTTCAGATCAGAACTTCGGAAGCTACTTTTCCCAATCAAC
120 130 140 150 160 170 180
AACACCCCTTCAGGATTTAAACCAAGGGGACACTGGATCACCTAGTGTTCACAGCAGGTAACCTCTG
|||||
GGGACACTGGATCACTAGTGTTCACAGCAGGTAACCTCTG
X 10 20 30 40
190 200 210 220 230 240 250
CTGTAGGAGAGAGAACTAAGTCTGAAAGACCTGTGCTTTTCCACCAGGAAGTTTACTTGGGCATCTCC
|||||
CTGTAGGAGAGAGAACTAAGTCTGAAAGACCTGTGCTTTTCCACCAGGAAGTTTACTTGGGCATCTCC
50 60 70 80 90 100 110
260 270 280 290 300 310 320
TGAGCTAGGCAATAGCTGTAGGTGACTTCTGGAGCCTATCCCTTCCCGCCCGCCCAAAAGAGGGAG
|||||
TGAGCTAGGCAATAGCTGTAGGTGACTTCTGGAGCCTATCCCGCTTCCCGCCCGCCCAAAAGAGGGAG
120 130 140 150 160 170 180
330 340 350 360 370 380 390
ATTTAACGGGGAGCTGGCGCCAGAGCTGGGGAATGGCCCGGAGCCAGCGGCTTCTTCTCTCTGATG
|||||
ATTTAACGGGGAGCTGGCGCCAGAGCTGGGGAATGGCCCGGAGCCAGCGGCTTCTTCTCTCTGATG
```


[illegible]

```
Initial score      = 516  Optimized score = 516  Significance = -0.07
Residue Identity  = 99%   Matches         = 516  Mismatches  = 1
Gaps              = 0     Conservative Substitutions = 0
```

5410 5420 5430 5440 5450 5460 5470
ACAAGCTGACTGCTCTCCAAAGTGACACTGTTTAGAGTCCAATCTTAGGCACACAAATGGTGCTCTCT
CCTGTAGCTGTTTTCGTGAAAGGGTATTTCTCTCTCCACCTATATAGAGGAAGTGAAGTTCACAGTGT
|||||
TATTTCTCTCTCCAACTATATAGAAGGAAGTGAAGTTCAGTCT
X 10 20 30 40
5480 5490 5500 5510 5520 5530 5540 5550
TTCTGGCAAGGGTAAACAGATCCCTCTCTCATCTCTTCTCTCAAGTGCCTCTTTGGTGAAG
50 60 70 80 90 100 110
TTCTGGCAAGGGTAAACAGATCCCTCTCTCTCTCTCTCTCTCTCAAGTGCCTCTTTGGTGAAG
530 5640 5650 5660 5670 5680 5690
GTGACATCATGTGACCTCTTCAGTGACCACTTACCGGTCTGGGGCTTGAACCTACCCCCAGACAATCT
GTGACATCATGTGACCTCTTCAGTGACCACTTACCGGTCTGGGGCTTGAACCTACTACCCCCAGACAATCT

Initial Score	=	38	Optimized Score	=	38	Significance	=	-0.26
Residue Identity	=	97%	Matches	=	38	Mismatches	=	1
Gaps	=	0	Conservative Substitutions	=	0			0
<hr/>								
5720	5730	5740	5750	5760	5770	5780		
ATAAGCAGCCAATGGATGCCAAGAGGAGTTCGAACCTAAAGACGTAATTGCCCAATGGGATGGACCTACCAGG								
5790	5800	5810	X	5820	5830	5840	5850	X
GCTGGATAACCTTGGCTGTACCCCTCGGGAAGAGCAGAGATATACGTNCCAGGTGGAGACACCCAGGCGCTGG								
GGAAGAGCAGAGATATACGTACCAGGTGGAGCACCAGG								
X 10 20 30 X								
<hr/>								
5860	5870	5880	5890	5900	5910	5920		
ATCAGCCCTCATTTGTGATCTGGGGTATGTGACTGATGAGACGAGGACCTGAGAAATCTATTGGGGGTTG								
5930	5940	5950						
AGAGGAGTGCCTGAGGAGGTAAT								
<hr/>								
14. US-09-497-957-1 (1-10825)								
US-09-497-957-29 Sequence 29, Application US/09497957								
Initial Score	=	38	Optimized Score	=	38	Significance	=	-0.26
Residue Identity	=	97%	Matches	=	38	Mismatches	=	1
Gaps	=	0	Conservative Substitutions	=	0			0
<hr/>								
5720	5730	5740	5750	5760	5770	5780		
ATAAGCAGCCAATGGATGCCAAGAGGATTCGAACCTAAAGACGTAATTGCCCAATGGGATGGACCTACCAGG								
5790	5800	5810	X	5820	5830	5840	5850	X
GCTGGATAACCTTGGCTGTACCCCTCGGGAAGAGCAGAGATATACGTNCCAGGTGGAGACACCCAGGCGCTGG								
GGAAGAGCAGAGATATACGTGCCAGGTGGAGCACCAGG								
X 10 20 30 X								
<hr/>								
5860	5870	5880	5890	5900	5910	5920		
ATCAGCCCTCATTTGTGATCTGGGGTATGTGACTGATGAGACGAGGACCTGAGAAATCTATTGGGGGTTG								
5930	5940	5950						
AGAGGAGTGCCTGAGGAGGTAAT								
<hr/>								
15. US-09-497-957-1 (1-10825)								
US-09-497-957-43 Sequence 43, Application US/09497957								
Initial Score	=	31	Optimized Score	=	31	Significance	=	-0.26
Residue Identity	=	96%	Matches	=	31	Mismatches	=	1
Gaps	=	0	Conservative Substitutions	=	0			0
<hr/>								
3770	3780	3790	3800	3810	3820	3830		
GGTTACACTCTCTGCACCTACTCTTATGGGTGCCTCAGACGAGGACCTTGGCTTTCTCTTTGTTGAAGCT								
3840	3850	3860	3870	3880	3890	X	3900	
TTGGGCTACGTGGATGACCACTGTTCGCTTCATGATNATCAGATCGCGCTGTGGAGCCCCCACTCA								
TGTTCTATGATCATGAGAGTCGCCGTGTGGAG								
X 10 20 30 X								
<hr/>								
3910	3920	3930	3940	3950	3960	3970		
TGGGTTTCCAGTAGAATTTCAAGCCAGATGTGGTCGAGTCAGAGTCTGAAGGGTGGGATCACATG								
3980	3990							
TTCACGTGTTGACTCT								
<hr/>								
16. US-09-497-957-1 (1-10825)								
US-09-497-957-44 Sequence 44, Application US/09497957								
Initial Score	=	30	Optimized Score	=	30	Significance	=	-0.26
Residue Identity	=	93%	Matches	=	30	Mismatches	=	2
Gaps	=	0	Conservative Substitutions	=	0			0

3930 3940 3950 3960 3970 3980 3990
CCAGATGTGGCTGACGTGACTCAGACTCTGAAGGGTGGGATCACATCTTCACTTGTGACATTCGGACATAT
|||||
CTGAAGGGTGGGATCAGAT
X 10 20

4000 4010 4020 4030 4040 4050 4060 4070
TATGGAAATCACAACCAACAGCAAGGTATGTGGAGAGGGGGCCCTACCTTCCCTGAGGTTGTCAAGACTTTT

CATC

20. US-09-497-957-1 (1-10825)

US-09-497-957-13 Sequence 13, Application US/09497957

Initial Score	=	20	Optimized Score	=	20	Significance	=	-0.26
Residue Identity	=	100%	Matches	=	20	Mismatches	=	0
Gaps	=	0	Conservative Substitutions	=	0		=	0

5460 5470 5480 5490 5500 5510 5520
CTTAGGACACAAAAATGGTGTCCTCTGTAGCTGTCTTTCTGAAAAGGGTATTTCTTCTCTCAACCTA

5530 5540 5550 5560 5570 5580 5590
TAGAAGGAGTGAAGTTCGAGTCTTCTGGCAAGGTAACAGATGCCCTCTCTCTATCTCTCTCTCTTCC
|||||
TGGCAAGGTAACAGATCC
X 10 20

5600 5610 5620 5630 5640 5650 5660 5670
TGTC AAGTGCCTCTTTGGTGAAGGTGACACATCATGTGACCTCTTCAGTGACCACTCTACGGTGTGCGGGC

TTGA

